Influence of front-of-package nutrition labels on beverage healthiness perceptions: Results from a randomized experiment

Rachel B. Acton a, Lana Vanderlee b, David Hammond a,⁎, 1

a School of Public Health and Health Systems, University of Waterloo, Waterloo, ON, Canada
b Department of Nutritional Sciences, University of Toronto, Toronto, ON, Canada

Abstract

The current study explored the influence of three summary indicator front-of-pack (FOP) nutrition labels on consumer perceptions of the healthiness of different beverage products. In 2016, a total of 675 respondents in southwestern Ontario aged 16 and over viewed images of soda, unflavoured milk and chocolate milk displaying one of four FOP label conditions (no FOP label, numeric rating, health star rating (HSR), or simplified traffic light (STL)), and rated the products’ healthiness. Participants also indicated their preference for summary indicator versus nutrient-specific FOP labels. Logistic regression models comparing correct responses across label conditions found no differences across label conditions for unflavoured milk or soda. Consumers in the HSR and STL conditions were more likely to correctly perceive a chocolate milk beverage as ‘moderately healthy’ (p = 0.004, p = 0.016). No differences in responses were identified across sociodemographic groups. Most respondents (93%) indicated that they would like to see a health rating or nutrient-specific information on the front of food products. Results of this study suggest that the influence of FOP labels may vary based on the nutritional quality of food products, and may have the greatest influence on consumer perceptions of ‘nutritionally ambiguous’ foods. Consumers indicated almost unanimous support for implementing FOP nutrition labelling systems.

1. Introduction

Poor diet is an important risk factor for numerous chronic diseases (World Health Organization, Food and Agriculture Organization of the United States, 2003). Recent shifts towards predominantly processed and calorie-dense diets have led to increasing rates of overweight and obesity in both high-income countries and low- and middle-income countries (World Health Organization, 2016; Popkin, 2001). In 2014, it was estimated that the global economic impact of obesity was approximately $2.0 trillion per year (McKinsey Global Institute, 2014).

Nutrition labelling is an intervention that has been used to support healthier dietary intakes at the population level (Roberto and Khandpur, 2014). In Canada, as in most other countries, food manufacturers are required to disclose the nutrient information of their foods via tables or panels displayed on the back of their food packages (Government of Canada, 2015). Many consumers report using these nutrition panels (Cowburn and Stockley, 2005; Canadian Council of Food and Nutrition, 2008); however, the nutrition information presented on the back of packages requires a high level of health literacy, and many consumers struggle to interpret the quantitative information, particularly with respect to serving size and percent daily values (Cowburn and Stockley, 2005; Acton et al., 2016; Hobin et al., 2016; Vanderlee et al., 2015).

Simplified ‘interpretive’ nutrition labels have emerged as an important intervention to complement the quantitative information provided on back-of-package nutrition panels. A range of interpretive front-of-package (FOP) labelling systems have been implemented internationally, most of which can be categorized as either ‘nutrient-specific’ or ‘summary indicator’ labelling systems. Nutrient-specific systems display the amounts of specific nutrients, often within the context of other information, such as percent daily values, colour-coded traffic light systems, or interpretive text indicating the recommended frequency of consumption (UK Department of Health, 2016; Freire et al., 2016; Center for Science in the Public Interest, 2015). Nutrient-specific systems typically highlight ‘negative’ nutrients such as sugar, sodium, or saturated fat, although some systems also highlight ‘positive’ nutrients such as protein or calcium. In contrast, summary indicator systems assign one overall health rating to products based on

⁎ Corresponding author.
E-mail addresses: rbacton@uwaterloo.ca (R.B. Acton), dhammond@uwaterloo.ca (D. Hammond).
1 Contact info: University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada.

https://doi.org/10.1016/j.ypmed.2018.08.022
Received 27 March 2018; Received in revised form 10 August 2018; Accepted 21 August 2018
Available online 23 August 2018
0091-7435/ © 2018 Elsevier Inc. All rights reserved.
algorithms that consider the entire nutrient profile of a product. Summary indicators typically communicate the overall dietary quality of products using a single numeric scale (e.g., 1–100), or via three- or five-level star ratings (NuVal, 2017; Australian Government Department of Health and Ageing, 2016; Guiding Stars, 2017; Olstad et al., 2015). Although colour-coded traffic light systems have most commonly been used in nutrient-specific systems, simplified traffic light systems have also been tested for use in summary indicator labels (UK Food Standards Agency, 2004).

An increasing number of countries have implemented, or are developing, FOP labelling systems; however, there is no consensus on whether nutrient-specific or summary indicator systems represent best practice. Summary indicator systems have been implemented in countries such as Australia and New Zealand (the ‘Health Star Rating’), while France has developed the 5-Colour Nutri-Score system (Australian Government Department of Health and Ageing, 2016; World Health Organization, 2017). In contrast, countries such as Ecuador, Chile and the UK have opted for nutrient-specific formats, as has Health Canada in their proposed FOP labelling system (UK Department of Health, 2016; Freire et al., 2016; Center for Science in the Public Interest, 2015; Health Canada, 2016).

Several studies have compared the impact of nutrient-specific versus summary indicator labels on consumer perceptions and food choices (Hersey et al., 2013; Hawley et al., 2013; Cecchini and Warin, 2016); however, there is less evidence comparing the relative effectiveness of different summary indicators (Emrich et al., 2014; Méjean et al., 2014; Savoie et al., 2013; Feunekes et al., 2008; Ducrot et al., 2015). Research to date has suggested that summary indicators are generally effective at communicating the healthiness of products to consumers (Watson et al., 2014; Talati et al., 2017), but evidence comparing the relative effectiveness of different summary indicator formats is lacking. In addition, few studies have explored potential differences in consumers’ interpretation of FOP labels across different sociodemographic groups such as age, gender or ethnicity.

Although most labelling systems target all packaged food and beverage products, beverages provide a useful canvas for comparing FOP labelling systems across products. Beverage products, particularly sugary drinks, are the focus of much recent public health attention (World Health Organization, 2018; Jones et al., 2017; Centers for Disease Control and Prevention, 2017). Sugary drinks contribute a substantial volume of caloric energy due to their high sugar content; they provide little to no nutritional value and are associated with poorer dietary quality; and they lead to lower feelings of satiety compared to foods, resulting in higher energy intake overall (Malik et al., 2010; Malik et al., 2006). In addition, beverages have simple nutrient profiles relative to foods and are easily recognizable by most consumers regardless of brand or variety, making them a useful product category for comparisons.

The current study sought to investigate the relative impact of three summary indicator formats on consumers’ interpretation of the ‘healthiness’ of beverage products: numeric, star, and simplified traffic light labels. The study also examined potential differences in efficacy across sociodemographic groups, as well as consumer preferences between nutrient-specific and summary indicator labels. Results from this study have the potential to guide and inform labelling policy in jurisdictions that are considering FOP summary indicator systems.

2. Methods

The study was conducted in September and October 2016 as a component of a larger study (Acton and Hammond, 2018). Ethical clearance was received from the Office of Research Ethics at the University of Waterloo. Written consent was obtained from all participants.

2.1. Participants

Participants aged 16 years and older were recruited using convenience sampling in a shopping mall in southwestern Ontario, Canada. Canadian research ethics guidelines do not require parental consent for individuals aged 16 years or older. Other than age, no other exclusion criteria were used in order to provide a heterogeneous sample of consumers. Research assistants approached potential participants to ask whether they were interested in participating in a “study on beverage purchasing”, and provided no additional information on the nature of the study’s research questions. A total of 686 participants completed the study (452 refusals, 14 incompletes); 11 participants were removed due to data quality concerns (e.g., due to significant cognitive difficulties, visual impairment, or visible influence from peers), resulting in a final analytic sample of 675.

2.2. Protocol

The study consisted of two components: 1) a between-group experiment examining the influence of various summary indicator labels and 2) a question administered after the experiment examining participants’ preferences between a summary indicator and a nutrient-specific FOP label.

2.2.1. Summary indicator experiment

Participants were randomly assigned to one of four label conditions: no FOP labelling (control), a numeric rating (1–100), a health star rating (0.5–5 stars), or a simplified traffic light symbol (red, yellow, green). The FOP labell design and product scoring of the numeric rating were based on existing and proposed nutrition rating systems. The label design and product scoring of the numeric rating were based on the NuVal shelf price tag labelling system developed in the United States (NuVal, 2017), the health star ratings were based on Australia and New Zealand’s Health Star Rating System (Australian Government Department of Health and Ageing, 2016), and the traffic light symbol was designed based on a simplified version of the traffic light system used in the United Kingdom (UK Department of Health, 2016; UK Food Standards Agency, 2004). All three rating systems take into account the nutritional profile of a food or beverage product when assigning scores.

Within their assigned label condition, participants were shown three beverages: a 473 mL carton of Neilson unflavored 2% milk (‘healthy’), a 473 mL carton of Neilson 1% chocolate milk (‘moderately healthy’), and a 591 mL bottle of generic soda (‘unhealthy’), all of which displayed the FOP label corresponding to the label condition that the participant was assigned to. The generic brand of soda was created for the current study to eliminate any pre-existing brand perceptions; the Neilson brand was selected for the milk products due to the brand’s high availability in the region, and because there is little variation in the nutrient profiles across other brands of unflavoured and chocolate milks. The order in which the three beverages were presented to participants was randomized. Participants viewed the ‘front’ of each product, as well as an image of the ‘back’ of each product, which featured the Nutrition Facts table (NFT), required on all pre-packaged foods in Canada. Therefore, participants were presented with information on specific nutrient amounts for each product. Nutrient profiles reflected those of commercially available products for Neilson brand 2% unflavoured milk and 1% chocolate milk. The generic soda beverage was assigned a nutrient profile reflecting that of a regular/non-diet lemon-lime sports drink (Gatorade brand); this lower sugar profile was selected to ensure that the sugar content was not overly obvious to respondents. Images of the beverages with each label condition and their corresponding NFTs are provided in Fig. 1.

Participants rated each of the three beverages using the same measure: “Would you consider this product to be...”, with the response options ‘unhealthy’, ‘moderately healthy’ and ‘healthy’. ‘Don’t know’
and ‘refuse to answer’ were also available as valid response options for each question. Responses for the unflavoured milk, chocolate milk and soda were coded as ‘correct’ based on responses of ‘healthy’, ‘moderately healthy’, and ‘unhealthy’, respectively.

2.2.2. Preference for summary indicator versus nutrient-specific FOP labelling

Upon conclusion of the FOP labelling experiment, all participants were asked, “What information would you most like to see on the front of food and beverage products?”, with the response options: ‘a rating of how healthy the product is overall’, ‘warnings for high levels of specific nutrients, such as sugar or sodium’, ‘both of the above’, ‘neither of the above’, ‘something else’, ‘don’t know’, or ‘refuse to answer’. Sample images were provided corresponding to the first two response options: a health star rating label was provided as an example of an overall rating, and a symbol denoting ‘high sugar’ was provided as an example of a nutrient-specific label (see Fig. 2).

2.2.3. Sociodemographic measures

Participants reported their age, gender, ethnicity, height and weight. Age was collected as an open-ended continuous variable, but has been categorized in the sample characteristics table. Self-reported height and weight were used to calculate body mass index (BMI), which was categorized into “underweight”, “normal weight”, “overweight” and “obese” using the World Health Organization (WHO) thresholds (World Health Organization, n.d.-a). BMI for participants 19 years of age or younger were calculated using BMI-for-age percentile growth charts as recommended by the Centre for Disease Control and Prevention and WHO guidelines (World Health Organization, n.d.-b; Centers for Disease Control and Prevention, n.d.).

2.3. Analysis

Statistical analyses were conducted using SPSS software (version 24.0; IBM Corp., Armonk, NY; 2015). Chi-square tests were used to test
for sociodemographic differences between experimental conditions to ensure successful randomization in the between-group experiment. Descriptive statistics were used to identify participant responses in each task by label condition. To examine the effect of label condition on correct responses, binary logistic regression models were fitted for each of the three beverage tasks, where 1 = correct, and 0 = incorrect/don’t know. All three models included a predictor variable for label condition (control, numeric, star rating, traffic light). Dummy variable coding was used to explore all possible contrasts between the four label condition categories. To explore potential moderating effects of the sociodemographic characteristics, additional binary regression models were fitted for each of the three outcomes, with predictor variables for age (continuous, 16–82), gender (male, female), ethnicity (white, non-white, indigenous) and BMI (underweight, normal weight, overweight, obese, not reported) in addition to label condition. Interaction terms between label condition and each of the sociodemographic variables were included in the models to test for moderation. Descriptive statistics were used to examine participants’ preferences between summary indicator and nutrient-specific FOP nutrition labels. The significance threshold was set at 0.05 for all tests.

3. Results

Sample characteristics are presented in Table 1. There were no significant differences in sociodemographic measures between the experimental conditions assigned in the between-group experiment, indicating that randomization was successful.

3.1. Impact of FOP label conditions

Table 2 shows responses across all experimental conditions for each of the three beverage tasks. Overall, a majority of participants across all conditions perceived unflavoured milk to be “healthy”, soda to be “unhealthy”, and chocolate milk to be “moderately healthy”.

Fig. 3 illustrates the proportion of participants who reported the “correct” response for each beverage by experimental condition. There were no significant differences in the proportion of respondents who correctly perceived the unflavoured milk as “healthy” between any of the conditions (see Fig. 3a) (all p > 0.05).

In contrast, for chocolate milk, participants who viewed beverages with a health star rating were significantly more likely to correctly identify chocolate milk as ‘moderately healthy’ than participants who saw no enhanced FOP labelling (see Fig. 3b) (OR = 1.96, 95% CI 1.24–3.09, p = 0.004). This was also true for participants who viewed beverages with a simplified traffic light symbol compared to no enhanced label (OR = 1.70, 95% CI 1.11–2.63, p = 0.02). However, participants who viewed the numeric rating were no more likely to correctly identify chocolate milk as ‘moderately healthy’ than those in the control condition (OR = 1.38, 95% CI 0.89–2.13, p = 0.15). There were also no significant differences when comparing the numeric rating to the star rating (OR = 1.42, 95% CI 0.88–2.28, p = 0.15) or traffic light symbol (OR = 1.24, 95% CI 0.79–1.94, p = 0.36). The health star and traffic light groups also did not significantly differ from one another (OR = 0.871, 95% CI 0.54–1.40, p = 0.57).

Finally, we examined the proportion of participants in each label condition who correctly identified the soda beverage as ‘unhealthy’. Similarly to unflavoured milk, there were no significant differences in correct responses between any of the conditions (see Fig. 3c) (all p > 0.05).

3.2. Sociodemographic influences on correct responses

None of the interactions between label condition and the sociodemographic variables age, gender, ethnicity, or BMI revealed significant moderating effects on correct responses for any of the three beverage tasks (all p > 0.1).

3.3. FOP label format preference

When asked which form of nutrition information they would most like to see on the front of food and beverage products, the largest proportion of participants (45%) indicated that they would like to see both an overall health rating and nutrient-specific information. Similar numbers of participants indicated a preference for either an overall health rating (24%) or nutrient-specific labels (25%), while 4% of participants selected ‘neither of the above’, 2% indicated ‘something else’, and 1% selected ‘don’t know’.

4. Discussion

The findings suggest that summary indicator labels were most impactful when communicating the nutrition level of a ‘moderately’ or ‘ambiguously’ healthy beverage product. None of the labels tested had a significant impact on beverage perceptions for products people already

---

**Table 1** Sample sociodemographic characteristics overall (n = 675) and by label condition.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total sample</th>
<th>Control (no label)</th>
<th>Numeric rating</th>
<th>Star rating</th>
<th>Traffic light</th>
<th>Chi square tests of independence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16–18</td>
<td>15.1% (102)</td>
<td>14.4% (26)</td>
<td>13.5% (22)</td>
<td>17.5% (27)</td>
<td>15.2% (27)</td>
<td>8.2 (0.77)</td>
</tr>
<tr>
<td>19–21</td>
<td>27.3% (184)</td>
<td>25.0% (45)</td>
<td>26.4% (43)</td>
<td>26.6% (41)</td>
<td>30.9% (55)</td>
<td></td>
</tr>
<tr>
<td>22–24</td>
<td>13.8% (93)</td>
<td>17.8% (32)</td>
<td>16.0% (26)</td>
<td>11.0% (17)</td>
<td>10.1% (18)</td>
<td></td>
</tr>
<tr>
<td>25–45</td>
<td>25.0% (169)</td>
<td>23.9% (43)</td>
<td>27.0% (44)</td>
<td>25.3% (39)</td>
<td>24.2% (43)</td>
<td></td>
</tr>
<tr>
<td>46+</td>
<td>18.8% (127)</td>
<td>18.9% (34)</td>
<td>17.2% (28)</td>
<td>19.5% (30)</td>
<td>19.7% (35)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9 (0.84)</td>
</tr>
<tr>
<td>Male</td>
<td>46.1% (311)</td>
<td>46.7% (84)</td>
<td>47.2% (77)</td>
<td>42.9% (66)</td>
<td>47.2% (84)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>53.9% (364)</td>
<td>53.3% (96)</td>
<td>52.8% (86)</td>
<td>57.1% (88)</td>
<td>52.8% (94)</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.6 (0.36)</td>
</tr>
<tr>
<td>White</td>
<td>52.6% (355)</td>
<td>57.2% (103)</td>
<td>49.1% (80)</td>
<td>55.2% (85)</td>
<td>48.9% (87)</td>
<td></td>
</tr>
<tr>
<td>Non-white/not reported</td>
<td>44.9% (303)</td>
<td>40.0% (72)</td>
<td>49.7% (81)</td>
<td>41.6% (64)</td>
<td>48.3% (86)</td>
<td></td>
</tr>
<tr>
<td>Indigenous</td>
<td>2.5% (17)</td>
<td>2.8% (5)</td>
<td>1.2% (2)</td>
<td>3.2% (5)</td>
<td>2.8% (5)</td>
<td></td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.8 (0.80)</td>
</tr>
<tr>
<td>Underweight (&lt; 18.5)</td>
<td>4.1% (28)</td>
<td>5.0% (9)</td>
<td>3.7% (6)</td>
<td>1.9% (3)</td>
<td>5.6% (10)</td>
<td></td>
</tr>
<tr>
<td>Normal weight (18.5–24.9)</td>
<td>47.9% (323)</td>
<td>47.2% (85)</td>
<td>45.4% (74)</td>
<td>50.6% (78)</td>
<td>48.3% (86)</td>
<td></td>
</tr>
<tr>
<td>Overweight (25.0–29.9)</td>
<td>23.3% (157)</td>
<td>25.6% (46)</td>
<td>27.0% (44)</td>
<td>22.1% (34)</td>
<td>18.5% (33)</td>
<td></td>
</tr>
<tr>
<td>Obese (30+)</td>
<td>14.8% (100)</td>
<td>13.3% (24)</td>
<td>14.7% (24)</td>
<td>14.9% (23)</td>
<td>16.3% (29)</td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>9.9% (67)</td>
<td>8.9% (16)</td>
<td>9.2% (15)</td>
<td>10.4% (16)</td>
<td>11.2% (20)</td>
<td></td>
</tr>
</tbody>
</table>
Information for the beverages throughout the summary indicator experiments. Limitations include the use of a non-probability based sample, meaning that the sample may not be representative of the Canadian population. In addition, although the images used in the experiments were designed to depict actual beverages as accurately as possible, the use of laptop-based tasks may not represent how consumers naturally interact with labels in a real-world setting. We also only tested three beverage products and do not know how such labels might influence perceptions of a wider range of food and beverage products. Further, we did not test beverage packages with any additional health claims or marketing; it is possible that FOP labels have less influence when competing with well-known branding and other marketing on packaging. Finally, our moderate sample size likely limited our power to test for interactions and we did not measure moderators such as numeracy or literacy that may be more predictive of label response.

5. Conclusions

The findings suggest that FOP summary indicator labels may be most impactful when communicating the healthiness of beverage products with ‘moderately’ or ‘ambiguously’ healthy nutrient profiles. In addition, more intuitive summary indicator labels, such as a star rating or traffic light symbol, showed the largest difference in outcomes in comparison to no FOP labelling. These results reemphasize for researchers and practitioners that the average consumer is of course not ignorant of basic nutrition knowledge; rather, consumers are most likely to find FOP labelling systems useful in cases when a product’s nutrition profile, or ‘healthiness’, is less obvious. These experimental findings provide evidence that may help guide policy decisions in jurisdictions where a summary FOP nutrition labelling system is being considered. Future research should examine differences between interpretive labelling systems in their dissuasive effects for unhealthy foods, and whether these differences are similar to the impact on promoting healthier alternatives.

Conflict of interest

David Hammond has provided paid expert testimony on behalf of public health authorities in response to legal challenges from the food and beverage industry.

Acknowledgments

This study was supported by a Chair in Applied Public Health funded by the Public Health Agency of Canada in partnership with the
We thank the participants who contributed their time to the study, as well as the research assistants who assisted with participant recruitment and survey administration.

References


