A Brunswik lens model of consumer health judgments of packaged foods

JACOB L. ORQUIN*
Department of Business Administration/MAPP, Aarhus University, Bartholins Allé 10, Aarhus, 8000, Denmark

ABSTRACT

Consumer health judgments of packaged food were compared with an objective healthfulness criterion using a Brunswik lens model. Consumer judgments were obtained from a representative consumer sample (N = 1329) who evaluated the healthfulness of 198 packaged food products. The objective healthfulness criterion was calculated for each product according to its specific nutrition values using a validated nutrition profile. The lens model included explicit cues such as nutrition values, nutrition and health claims, food category, and brand and implicit cues such as a packaging design and category representativeness. The study revealed that the objective healthfulness criterion is highly predictive on the basis of cues such as the food category, brand, carbohydrate content, and whether the food is a typical “light” product. However, consumer judgments of food healthfulness are based almost entirely on the food category and to a lesser extent on the brand and consumer familiarity with the product. The results are in conflict with consumers’ self-reported use of nutrition information but are in accordance with findings from studies using implicit methods. Copyright © 2014 John Wiley & Sons, Ltd.

CONSUMER JUDGMENTS OF FOOD HEALTHFULNESS

Healthy eating has become an integral part of consumer policy in Europe and represents a latent conflict between politicians, consumers, and the industry. Although healthy eating is a multifaceted topic, most of the discussions seem to concern the nutritional aspects of food such as how to reduce the intake of saturated fats and sugars and increase the intake of fiber and protein (Chrysochou et al., 2010). A common policy approach has been to communicate the healthfulness of foods using nutrition labels, nutrition and health claims, or front-of-pack nutrition labels (Cowburn and Stockley, 2005). The underlying assumption is that adequate nutrition labeling empowers the consumer to judge the healthfulness of food products and thereby choose a healthier product. However, given the information complexity of nutrition labels, even the simplest strategy for processing the label requires several cognitive operations (for an overview, see Schulte-Mecklenbeck et al., 2013). First of all, the label information must attract sufficient visual attention to enter working memory where it must lead to an adjustment in the representation of the food nutrients (Graham et al., 2012). Second, given an adequate representation of the nutrient values, the consumer must now compare the product under consideration with a competing product or with a threshold value for the important nutrients. Finally, the consumer must choose the product with the most favorable nutrition values or in the case of a single product accept or reject it based on whether it is above the threshold value (Schulte-Mecklenbeck et al., 2013). If the healthfulness judgment is to adequately reflect the intention behind the nutrition label, one could furthermore add that the judgment must be compensatory, that is, the consumer must take into account more than one nutrition value and if necessary trade off between these values. The number of cognitive operations is typically higher for compensatory compared with noncompensatory processing strategies (Gigerenzer and Gaissmaier, 2011), which suggests that making healthfulness judgments from nutrition labels is a rather effortful process (Johnson and Payne, 1985). The effort associated with reading nutrition labels should be familiar to anyone who has ever tried to compare the nutrition values of two or more products in the supermarket. The question is, of course, whether consumers can be expected to go through such an extensive process.

Prior research suggests that consumers rely on “fast and frugal” decision strategies for food choice (Scheibehenne et al., 2007; Schulte-Mecklenbeck et al., 2013) as well as simplistic and categorical thinking about food and nutrition (Rozin et al., 1996; Oakes and Slotterback, 2001a, 2001b; Chernev and Gal, 2010). These studies suggest that an effortful judgment process based on nutrition label information is not very likely to happen and that health judgments may be based on less demanding heuristics. These findings are furthermore corroborated by a group of studies examining consumer attention to product packaging using real-world stimuli. These studies have used eye movements as an objective and indirect measure of attention and examine what aspects of packaged foods are attended to in shopping or choice situations. The earliest of these studies was authored by Visschers et al. (2010) who examined consumer attention to breakfast cereal product packaging under health and hedonic taste choice goals. Visschers et al. found that consumers with a hedonic taste goal have the most fixations to the packaging design followed by the product name, text areas, and to a smaller extent nutrition tables and front-of-pack nutrition labels. Consumers with a health goal also have most fixations to the packaging design elements, but this group has significantly more fixations to the nutrition label compared with the taste goal group.

Later Bialkova and van Trijp (2011) studied consumer attention to yogurt product packaging under either a health or preference choice goal. Their results indicate that consumers with a preference goal spend most of their time gazing at the flavor information followed by the brand and nutrition information. Consumers with a health goal in mind spend most of their gaze time on nutrition information followed by flavor and brand information.
Orquin and Scholderer (2011a) examined the relative amount of fixations dedicated to product packaging for dairy products in different categories in health and purchase evaluation tasks. Comparing the health evaluation condition to the purchase evaluation condition revealed an increase in fixations to the nutrition table and fat percentage and a decrease in attention to the brand and design elements. The results also revealed a significant effect of health motivation, gender, and body mass index on fixations to the nutrition label.

Orquin et al. (2012) examined consumer attention to product packaging in a choice task involving more than 150 dairy products from different categories. The results indicate that consumers mainly gaze at the brand information followed by information about the product category and product images. Consumers also gaze at front-of-pack nutrition labels, organic labels, and health claims, yet to a much smaller extent.

Vidal et al. (2013) studied consumer attention to yogurt product packaging in a health judgment task and found that consumers had most fixations to the nutrition label, ingredient list, and front-of-pack nutrition label and to a much smaller extent the product image and brand elements.

Orquin et al. (2013) studied consumer attention to yogurt products in a discrete choice experiment. Consumers had most fixations to the flavor, fat content, and price information and to a smaller extent to the brand, organic label, and front-of-pack nutrition label.

Ares et al. (2014) segmented consumers into two clusters on the basis of deliberate versus intuitive thinking styles and examined the attention to yogurt product packaging for the two clusters. Consumers adhering to a deliberate thinking style had the most fixations to the nutrition label followed by the central product image and front-of-pack nutrition label. Consumers adhering to an intuitive thinking style had the most fixations to the central product image followed by the front-of-pack nutrition label and the nutrition label.

Overall, the eye-tracking studies reveal a high degree of consistency in their findings, which indicate that consumer attention to packaged food products may be limited to very few packaging elements. Consumers mostly process information about the brand, flavor, and product images, whereas nutrition-related information receives limited amounts of attention. The studies also show that health motivation has a large influence on attention to nutrition-related information and this is irrespective of whether health motivation is measured or manipulated experimentally. Although the conclusions from the eye-tracking studies corroborate a restricted or heuristic use of information in food choice (Scheibehenne et al., 2007; Schulte-Mecklenbeck et al., 2013), the findings are in conflict with studies using, for instance, self-assessment measures (Cowburn and Stockley, 2005) or verbal protocol analysis (Higgins et al., 2002b, 2002a), which generally indicate that consumers use nutrition and health-related information to a much greater extent. Although the eye-tracking studies may be more objective with regard to consumer attention to nutrition information, they also suffer from one main limitation: although attention is closely related to judgment and decision making (Orquin and Mueller Loose, 2013), there is no guarantee that attended information will influence the judgment.

The question about how consumers judge the healthfulness of packaged foods is by no means negligible as most policy interventions directed toward healthy food choice are based on a strong assumption about how consumers make such judgments. Given the ambiguity of the prior research and the importance to policy, this study examines consumer health judgments of packaged foods. To increase the external validity of the study and usefulness to policy makers, the study is based on real-world stimuli, that is, images of existing market products.

As a second and equally important goal, the study compares consumer health judgments with an objective healthfulness criterion based on the nutrition information for each individual food product. The latter question is particularly important as any mismatch between consumer health judgments and objective food healthfulness is an indicator that consumers are potentially being misled or misleading themselves in their food choice.

METHOD

Modeling consumer judgments

As has been described earlier, the literature reveals a conflict between two views regarding consumer judgments of food healthfulness. One view, entertained by policy makers and supported by research relying on self-report measures (Cowburn and Stockley, 2005), is that consumers process nutrition and health information in an extensive manner such as by reading nutrition labels and comparing nutrition values for different products. This view is in conflict with another group of studies suggesting that consumers rely on heuristic and categorical thinking about food healthfulness (Rozin et al., 1996) and eye-tracking studies suggesting that consumers devote limited attention to, for instance, nutrition label viewing (Graham et al., 2012).

First, given the ambiguity about how consumers judge food healthfulness, a sound methodological approach should be able to accommodate different classes of judgment processes. Second, because it is an explicit aim to measure health judgments of real-world product stimuli, further constraints are imposed as to the choice of methodology. Using real-world product stimuli excludes experimental approaches that rely on factor designs because any experimental manipulation to the stimuli means that the stimuli are not real anymore. Using real product stimuli therefore suggests the need for a quasi-experimental design, which means that the estimation approach should be able to handle an unbalanced design. One possibility is using a linear mixed model or a generalized linear mixed model, which has the advantage of being robust with regard to unbalanced designs, and this class of models can furthermore accommodate choice and judgment data drawn from different classes of decision strategies (Mueller Loose et al., 2012).

In the present study, the choice of modeling approach fell on the Brunswik lens model (Brunswik, 1952), which, in its modern application, is based on a linear model (Karelaia and Hogarth, 2008). This type of model compares the importance of several independent variables (called cues) and is useful...
for answering questions such as whether, and to what extent, consumers use nutrition information in judgments of food healthfulness. The model also provides a measure of judgmental accuracy in relation to the environment, which, in this case, means how accurately consumers judge the healthfulness of foods in comparison with an objective healthfulness criterion. This aspect of the model is particularly important in order to answer the question of whether consumers are being misled or misleading themselves in their health judgments of food products. The model not only provides a measure of judgmental accuracy but also describes what cues might have led to the particular judgmental accuracy.

The Brunswik lens model has been used to model human judgments in quite diverse areas and normally consists of two independent linear models using the same set of cues as predictors. The left side of the model describes the predictability of the criterion being judged, and the right side of the model describes the subjective judgment of the criterion (Figure 1). The lens model thus provides a measure of the cues utilized in a judgment, the ecological validity of the cues for predicting the criterion, and a measure of judgmental accuracy, called the achievement index, which is the correlation between the subjective judgment and the criterion (Hogarth and Karelaia, 2007). In most lens model studies, the weights are based on correlations between the individual cues, the judgment, and the criterion values. However, the present model contains a mix of continuous and multilevel categorical predictors and so necessitates an alternative approach. Instead of correlations, the present lens model therefore uses effect sizes as a gauge of cue utilization, similar to previous approaches based on analysis of variance (Hoffman et al., 1968; Schultz and Gustavson, 1978).

In the case of consumer health judgments of packaged food, the lens model provides a measure of the following: (1) the cues utilized by consumers for judging food healthfulness; (2) the ecological validity of these cues for predicting food healthfulness; and (3) an achievement index of judgmental accuracy. Using a lens model approach to consumer judgments should thus clarify central questions such as the relative importance of nutrition information compared with the importance of other cues, the consistency of consumer judgments, and the appropriateness of consumer cue utilization.

Study variables
To address the question of how consumers evaluate the healthfulness of packaged foods, the present study employed a quasi-experimental design using images of existing packaged foods as stimuli. In comparison with previous research, which has used food names (Oakes and Slotterback, 2001a, 2001b; Oakes, 2004), verbal descriptions of foods (Oakes and Slotterback, 2001a, 2001b; Carels et al., 2007), or a mix of verbal and pictorial representations of foods (Chernev and Gal, 2010; Irmak et al., 2011), this approach mimics the variation in nutrition values and packaging design given in a real-world consumer environment. The experimental stimuli comprised of 198 products in four main categories of foods (yogurt, cheese, milk, and butter), all available in Danish retail stores at the time of the study.1 The dependent variables consisted of measures of perceived healthfulness and product familiarity, which were obtained from consumers who rated each of the 198 products separately. The independent variables were coded for each product on dimensions such as nutrition information, nutrition score (Rayner et al., 2005; Scarborough et al., 2007, 2010), brand, holistic package design (Orth and Malkewitz, 2008), nutrition and health claims, food category, and category prototypicality (Loken and Ward, 1990; Loken, 2006).

The nutrition information was transcribed from the nutrition label for each individual product, and a nutrition score was computed using a validated nutrition profile model (for a description of how to compute the nutrition score, see Scarborough et al., 2007). The choice of coding for the holistic package design was motivated by results of Orth and Malkewitz’s (2008) work demonstrating that certain packaging designs can affect consumers’ perceptions of the

1The experimental stimuli included dairy products only as agreed to in the project consortium and project plans.
healthfulness of wines. The concept of holistic package design is best understood as a design style that integrates various design elements into a certain stylistic expression that conveys an image of the product as being, for instance, delicate or massive (examples of holistic package designs are found in APPENDIX A). Orth and Malkewitz found that a so-called natural design, characterized by the use of harmonious and elaborate design factors, was perceived as more healthful than other holistic designs, labeled as massive, contrasting, delicate, and nondescript. The validity of the holistic design approach has been demonstrated in different product categories including both foods and non-foods; furthermore, a holistic package design has been found to affect price expectations, attractiveness and quality perceptions (Orth et al., 2010), reaction times in product choice tasks (Reimann et al., 2010), and brand personality characteristics (Orth and Malkewitz, 2008).

Besides the holistic package design, all products were coded on prototypicality dimensions. The term prototypicality refers to how representative a category exemplar is for a specific category and is normally measured by letting participants rate category exemplars on a scale, with endpoints such as “good example/poor example” or “representative/unrepresentative” (Loken and Ward, 1990). The definition of prototypicality most commonly used in categorization research, and also the one adopted in this study for instructing the coders, stems from research by Rosch and Mervis (1975). In terms of the categorization of foods, increased prototypicality has been shown to increase liking for the product (Halberstadt, 2006; Förster and Denzler, 2011) as well as facilitate categorization of the exemplar product as belonging to a particular category (Bauer et al., 1995). A food’s category has previously been shown to play a prominent role in consumers’ judgments of a food’s healthfulness compared with, for instance, nutrition and health information (Orquin and Scholderer, 2011b). In terms of perceived healthfulness, this could mean that an increase in prototypicality might enhance the usefulness of the food category for judging healthfulness. If a food product has a high prototypicality, then it should be easier for consumers to categorize the product as belonging to a specific food category, which would make the category a more valid cue to the consumer for judging healthfulness. On the other hand, if a product has a low prototypicality, it should be more ambiguous what food category it belongs to, and the category should therefore be a less valid cue to the consumer judgment. To examine the effect of prototypicality, four measures were obtained: the first one being a taxonomic prototypicality measure and the remaining three being goal-oriented prototypicality measures relevant to the context of consumption such as supplement product prototypicality, diet product prototypicality, and snack product prototypicality. The prototypicality measures were obtained from an independent sample of coders.

Besides product-level characteristics, measures on consumer characteristics were obtained, which could be expected to affect perceptions of food healthfulness such as demographic variables, general health interest (Roininen et al., 1999), nutrition and health knowledge, and a subjective perception of own nutrition and health knowledge (Andrews et al., 2009).
could be due to the fact that the convenience sample consisted of participants naïve to product design. The holistic package categories were assigned to each product in the pilot test on the basis of a majority rule, and the assigned codes showed a fair agreement with the categories assigned in the Orth and Malkewitz study, Cohen’s $\kappa = 0.36$. The experimental stimuli were rated by 12 designers and revealed a fair inter-rater agreement, Cohen’s $\kappa = 0.32$. In cases with multiple modes, a new mode was calculated excluding the data from the least reliable coder.

Nutrition information and nutrition score
Nutrition information was coded for each product with regard to calories, fat, carbohydrates, and protein. The values were recorded per 100 g of the product following the Danish standardized guidelines (Ministry of Food, Agriculture and Fisheries, 2013). From these values, a score was computed for each product following the directions of the WXYfm nutrient profiling model (Scarborough et al., 2007, 2010). The WXYfm model is currently used in the UK in the regulation of food advertising, and a modified version of the model is used in Australia and New Zealand in the regulation of nutrition and health claims (Scarborough et al., 2010). The computation of the WXYfm model consists of two aggregate scores, one based on the content of negative nutrients such as saturated fat, sugar, calories, and sodium and the other score based on the content of positive nutrients such as dietary fiber, protein, fruits, vegetables, and nuts (Scarborough et al., 2007). The original computation of the WXYfm nutrition score contained an additional weighted penalty for high-fat foods, which was removed in the present study to produce a more continuous nutrition score. The nutrition score was furthermore reversed for ease of interpretation so that a more nutritious product would have a higher score than a less nutritious one.

Nutrition and health claims, keyhole label and organic label
Each product was coded (present or not present) for nutrition and health claims, organic labels, and keyhole labels. The stimulus set contained 14 products with nutrition and health claims (translated in APPENDIX B), 20 products with keyhole labels, and 42 products with organic labels. The coding of nutrition and health claims was based on the European Union regulation of nutrition and health claims made on foods, which states:

“nutrition claim” means any claim which states, suggests or implies that a food has particular beneficial nutritional properties (…).

“health claim” means any claim that states, suggests or implies that a relationship exists between a food category, a food or one of its constituents and health.

(Scarborough et al., 2007, 2010).

Category, supplement product, diet product, and snack product prototypicality
Prototypicality measures were adopted from Loken and Ward (1990) and measured exemplar goodness for individual category members according to a taxonomic categorization (category prototypicality) and three script categorizations (supplement product prototypicality, diet product prototypicality, and snack product prototypicality). The stimulus sample was blocked in five groups, and each block was rated by 12 coders on the four prototypicality measures using 10-point Likert scales following Rosch and Mervis (1975). The final data set showed high inter-rater agreement across the five blocks of coders (intraclass correlations ranging between 0.90 and 0.95). Because of large differences in scale use, standardized scores were computed at the coder level before computation of the mean prototypicality scores.

Dependent variables and consumer characteristics
The dependent variables consisted of two 7-point Likert items measuring product familiarity (1 = not at all familiar; 7 = highly familiar) and perceived healthfulness (1 = not at all healthy; 7 = very healthy). The perceived healthfulness item was adapted from Irmak et al. (2011). Additional information was collected on consumer demographics, a short scale measuring general health interest (Roizen et al., 1999), a single-item measure of subjective nutrition and health knowledge (“Compared to other people, how much do you feel you know about nutrition?”), and a six-item measure of nutrition and health knowledge (Andrews et al., 2009). Items for the nutrition and health knowledge questionnaire are reported in APPENDIX C.

Procedure
The first step in constructing the data set consisted of coding the stimuli with regard to the independent variables. The coding of holistic package design was carried out by designers ($N = 12$), each of whom coded all 198 products. For the coding of prototypicality, the stimuli were blocked in five groups and rated by a convenience sample ($N = 57$) in a between-subjects design. All products were coded on nutrition values as well as for nutrition and health claims and front-of-pack labeling. Step 2 of the data collection consisted of collecting the dependent variables from a consumer panel. Each participant in the consumer panel rated 22 products on perceived healthfulness.
and product familiarity. After completing the product ratings, the participants were given a questionnaire measuring consumer demographics, general health interest, and subjective and objective nutrition and health knowledge.

RESULTS

Analysis of consumer judgments
The first step in the analysis addressed consumer health judgments (the right side of the lens model) and used perceived healthfulness as a dependent variable. The analysis was carried out by means of a hierarchical linear model with observations as a Level 1 variable and participants as a Level 2 variable (Twisk, 2006). The final model included main effects of all product-level variables, an interaction effect of food category and category prototypicality, and product familiarity as independent variables. The tests of model effects are shown in Table 1 under the column “All consumers.” The final model had a pseudo-$R^2$ of 65.93 per cent, which indicates that consumers had a somewhat high degree of consistency in their judgments of food healthfulness (Karelaia and Hogarth, 2008). The second step in the analysis entailed creating two subsamples based on subjects’ nutrition and health knowledge (Appendix C). The group with the highest nutrition and health knowledge, dubbed nutrition experts, had either five or six correct answers out of six, whereas the group with the lowest nutrition and health knowledge, dubbed nutrition illiterates, had a score of either zero, one, or two correct answers. The two subsamples were analyzed using the same model specifications used for analyzing the total sample. Tests of model effects are shown in Table 1 under the columns “nutrition experts” and “nutrition illiterates.” The two models had a pseudo-$R^2$ of 67.73 per cent for the nutrition experts and a pseudo-$R^2$ of 67.24 per cent for the nutrition illiterates, indicating that the nutrition experts were only slightly more consistent in their judgments about the healthfulness of a product than the nutrition illiterates.

To visually examine the interaction effect between food category and category prototypicality, the slopes were plotted in Figure 2 on the basis of the analysis of the total sample. Interestingly, the slopes in Figure 2(A) seem to suggest that a higher prototypicality leads to a stronger effect of the category on perceived healthfulness, whereas the slopes in Figure 2(B) suggest the opposite, that is, a higher prototypicality diminishes differences between categories. The plots suggest that the effect of prototypicality on health judgments strongly depends on the particular food category although it is uncertain what aspect of the category drives the effect (see the Discussion section).

Analysis of the WXYfm model
The next step in the analysis addressed the ecological validity of the cues for predicting food healthfulness. The analysis was carried out by means of a linear mixed model using the WXYfm nutrition score as a dependent variable and all product-level variables as independent variables. Tests of model effects are shown in Table 1 under the column “WXYfm model.” The model had a pseudo-$R^2$ of 97.44 per cent, which means that the WXYfm nutrition score had a very high predictability on the basis of the independent variables. In other words, packaging and nutrition cues available to consumers have a high ecological validity.

The lens model
The lens model of consumer judgments and corresponding ecological validities of the cues were calculated using effect sizes as a means to gauge cue weights (Hoffman et al., 1968; Schultz and Gustavson, 1978). Effect sizes were calculated as the proportion reduction of residual variance for each model term by subtracting residual variance from the partial model without the term from the residual variance from the full model and dividing the product by the residual variance from the partial model (Peugh, 2010). The resulting lens model is presented in Table 2.

The achievement index, which is a measure of judgmental accuracy based on the correlation coefficient between subjective judgments and criterion values (Hogarth and Karelaia, 2007), resulted in a significant correlation between health judgments and the WXYfm score, $r = 0.338, p < 0.001$. Although significant, the achievement index is by no means high: a recent meta-analysis of lens model studies found that mean achievement levels are at 0.56 across different domains of human judgments (Karelaia and Hogarth, 2008). The achievement indices for the nutrition experts and the nutrition illiterates were 0.384 and 0.300, respectively.

The lens model revealed a strong effect of food category on both consumer judgments and the WXYfm model. The descriptive values in Table 3 show that consumer judgments are for the most part in the same direction as the WXYfm model; however, consumers are very inaccurate when it comes to judging the magnitude of healthfulness.

DISCUSSION

Summary of findings

Consumer heterogeneity
The study showed that consumers who have either high or low nutrition and health knowledge make almost the same judgments of food healthfulness, although they do differ on a couple of aspects. First of all, the model pseudo-$R^2$’s show that the high nutrition and health knowledge group is slightly more consistent in their evaluations than the low nutrition and health knowledge group (pseudo-$R^2$ of 67.73% vs. 67.24%). Second, for the high nutrition and health knowledge group, there is a smaller effect of familiarity than for the low nutrition and health knowledge group (proportion reduction in variance (PRV) of 0.018 vs. 0.065), meaning that the high-knowledge group is less biased by product familiarity in their judgments. Third, the high-knowledge group has a higher achievement index than the low-knowledge group ($r$ of 0.384 vs. 0.300), which indicates that they are better at judging the healthfulness of foods according to the WXYfm criterion. Apart from these differences, both groups were remarkably similar, with major effects relating to food category, familiarity, and

Copyright © 2014 John Wiley & Sons, Ltd.
Table 1. Tests of model effects and pseudo-$R^2$ for the full consumer judgment model, the nutrition experts, and nutrition illiterates (dependent variable: perceived healthfulness) and for the WXYfm model (dependent variable: WXYfm score)

<table>
<thead>
<tr>
<th>Factors</th>
<th>All consumers</th>
<th>Nutrition experts</th>
<th>Nutrition illiterates</th>
<th>WXYfm model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>num $df$</td>
<td>den $df$</td>
<td>$F$</td>
<td>$p$</td>
</tr>
<tr>
<td>Intercept</td>
<td>1</td>
<td>26,936.73</td>
<td>1031.58</td>
<td>0.00</td>
</tr>
<tr>
<td>Food category</td>
<td>11</td>
<td>24,901.88</td>
<td>37.26</td>
<td>0.00</td>
</tr>
<tr>
<td>Brand</td>
<td>28</td>
<td>25,089.69</td>
<td>17.79</td>
<td>0.00</td>
</tr>
<tr>
<td>Holistic package design</td>
<td>4</td>
<td>25,222.04</td>
<td>7.80</td>
<td>0.00</td>
</tr>
<tr>
<td>Nutrition/health claims</td>
<td>1</td>
<td>22,544.69</td>
<td>90.02</td>
<td>0.00</td>
</tr>
<tr>
<td>Keyhole label</td>
<td>1</td>
<td>24,924.47</td>
<td>4.91</td>
<td>0.03</td>
</tr>
<tr>
<td>Organic label</td>
<td>1</td>
<td>25,947.45</td>
<td>20.25</td>
<td>0.00</td>
</tr>
<tr>
<td>Category × prototypicality</td>
<td>11</td>
<td>25,798.36</td>
<td>6.23</td>
<td>0.00</td>
</tr>
<tr>
<td>Category prototypicality</td>
<td>1</td>
<td>26,261.03</td>
<td>4.19</td>
<td>0.04</td>
</tr>
<tr>
<td>Supplement prototypicality</td>
<td>1</td>
<td>25,457.04</td>
<td>28.75</td>
<td>0.00</td>
</tr>
<tr>
<td>Diet prototypicality</td>
<td>1</td>
<td>25,972.60</td>
<td>46.48</td>
<td>0.00</td>
</tr>
<tr>
<td>Snack prototypicality</td>
<td>1</td>
<td>26,371.08</td>
<td>40.75</td>
<td>0.00</td>
</tr>
<tr>
<td>Fat percentage</td>
<td>1</td>
<td>26,055.72</td>
<td>75.76</td>
<td>0.00</td>
</tr>
<tr>
<td>Carbohydrate percentage</td>
<td>1</td>
<td>21,303.67</td>
<td>1.13</td>
<td>0.29</td>
</tr>
<tr>
<td>Protein percentage</td>
<td>1</td>
<td>26,669.86</td>
<td>32.78</td>
<td>0.00</td>
</tr>
<tr>
<td>Calories</td>
<td>1</td>
<td>25,987.38</td>
<td>3.75</td>
<td>0.05</td>
</tr>
<tr>
<td>Familiarity</td>
<td>1</td>
<td>27,194.81</td>
<td>1674.43</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Note: num $df$=numerator degrees of freedom; den $df$=denominator degrees of freedom; pseudo-$R^2$ full consumer model=65.93 per cent; nutrition experts=67.73 per cent; nutrition illiterates=67.24 per cent; WXYfm model=97.25 per cent.
brand and minor or no effects for all other product-level characteristics. Because the differences between the two groups were minor, the following discussion of the findings refers to the full consumer judgment model.

Table 2. A lens model of information cues utilized in consumer judgments of food healthfulness and the corresponding ecological validities of the cues

<table>
<thead>
<tr>
<th>Factors</th>
<th>All consumers</th>
<th>Nutrition experts</th>
<th>Nutrition illiterates</th>
<th>WXYfm model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food category</td>
<td>0.209</td>
<td>0.207</td>
<td>0.251</td>
<td>0.232</td>
</tr>
<tr>
<td>Brand</td>
<td>0.031</td>
<td>0.033</td>
<td>0.032</td>
<td>0.366</td>
</tr>
<tr>
<td>Holistic package design</td>
<td>0.002</td>
<td>0.005</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>Nutrition/health claims</td>
<td>0.007</td>
<td>0.005</td>
<td>0.008</td>
<td>0.000</td>
</tr>
<tr>
<td>Keyhole label</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Organic label</td>
<td>0.010</td>
<td>0.010</td>
<td>0.009</td>
<td>0.011</td>
</tr>
<tr>
<td>Category prototypicality</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Supplement prototypicality</td>
<td>0.001</td>
<td>0.002</td>
<td>0.000</td>
<td>0.014</td>
</tr>
<tr>
<td>Diet prototypicality</td>
<td>0.002</td>
<td>0.003</td>
<td>0.002</td>
<td>0.123</td>
</tr>
<tr>
<td>Snack prototypicality</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>Category x prototypicality</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
<td>0.049</td>
</tr>
<tr>
<td>Fat percentage</td>
<td>0.002</td>
<td>0.004</td>
<td>0.002</td>
<td>0.005</td>
</tr>
<tr>
<td>Carbohydrate percentage</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.143</td>
</tr>
<tr>
<td>Protein percentage</td>
<td>0.003</td>
<td>0.001</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>Calories</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.039</td>
</tr>
<tr>
<td>Familiarity</td>
<td>0.039</td>
<td>0.018</td>
<td>0.065</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Note: Reported effect sizes are proportion reduction in variance (PRV). Note that the PRV values do not necessarily add up to 1.

Table 3. Descriptive statistics for the effect of food category on perceived healthfulness and WXYfm score

<table>
<thead>
<tr>
<th>Category</th>
<th>Consumer judgment (mean)</th>
<th>WXYfm model (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1% milk</td>
<td>0.644</td>
<td>1.075</td>
</tr>
<tr>
<td>0.5% milk</td>
<td>0.672</td>
<td>1.075</td>
</tr>
<tr>
<td>1.5% milk</td>
<td>0.290</td>
<td>0.883</td>
</tr>
<tr>
<td>3.5% milk</td>
<td>−0.099</td>
<td>0.379</td>
</tr>
<tr>
<td>Butter liquid</td>
<td>−0.510</td>
<td>−2.029</td>
</tr>
<tr>
<td>Butter hard</td>
<td>−0.378</td>
<td>−2.093</td>
</tr>
<tr>
<td>Cheese grated</td>
<td>−0.395</td>
<td>−0.473</td>
</tr>
<tr>
<td>Cheese block</td>
<td>−0.075</td>
<td>−0.344</td>
</tr>
<tr>
<td>Cheese sliced</td>
<td>−0.303</td>
<td>−0.374</td>
</tr>
<tr>
<td>Yogurt snack</td>
<td>0.045</td>
<td>0.568</td>
</tr>
<tr>
<td>Yogurt drink</td>
<td>−0.102</td>
<td>0.529</td>
</tr>
<tr>
<td>Yogurt 1-L pack</td>
<td>0.363</td>
<td>0.620</td>
</tr>
</tbody>
</table>

Note: Means are based on standardized values.

Food category and familiarity

The study revealed that consumer judgments of food healthfulness are to a large degree based on two factors: the food category and consumer familiarity with the individual product. With regard to the food category, consumer judgments were for the most part in the same direction as the WXYfm model, although the order of magnitude was seldom matched (Table 3). Moreover, consumers consistently underestimated the healthfulness of categories such as milk and yogurt and overestimated the healthfulness of categories such as butter and cheese.

With regard to the effect of familiarity, consumers generally perceived familiar products as more healthful than unfamiliar products. Such an effect probably arises because more familiar products are better liked, and this positive attitude spills over into other attitudes such as judgments of healthfulness. There can be little doubt, however, that consumers are misguided by this familiarity heuristic unless, of course, they have chosen to selectively expose themselves to healthful foods and thereby has become more familiar with healthful than with unhealthful foods.
A relevant question is whether category and familiarity effects function as heuristics in consumer health judgments of packaged foods. With regard to the category effect, it is worth noting that the effect size is in a completely different order of magnitude than any other cues in the model. Although the effect size by itself does not constitute proof of a heuristic, it does seem to corroborate previous literature showing that consumers generally use simplistic and categorical thinking about food and nutrition (Rozin et al., 1996; Oakes and Slotterback, 2001a, 2001b; Chernev and Gal, 2010). With regard to the familiarity effect on consumer judgments, this seems in accordance with a mere exposure effect, which predicts that the well-known is liked more than the less known (Zajonc, 1968; Moreland and Zajonc, 1977; Kunst-Wilson and Zajonc, 1980), which could spill over into perceptions of food healthfulness. Whether the familiarity effect is a consequence of mere exposure is not clear from the study, but it seems reasonable that the effect of familiarity is related to heuristic processing.

**Brand and holistic packaging design**

With regard to brand and packaging, the study showed a medium effect size of brand and a small effect size of holistic package design on consumer judgments. Interestingly, the brand had the largest effect size in the WXYfm model, probably because of the fact that some food producers position themselves on light-fat or reduced-fat products, whereas others produce only cheese or only butter. Regarding packaging, there was a small but significant effect of holistic package design; however, packaging design had no effect on the WXYfm model, which means that consumer perceptions of packaging (as communicating product healthfulness) do not reflect the actual healthfulness of foods. Such a situation is, of course, undesirable because it could indicate that consumers are misled by packaging designs to believe that some products are more healthful than they actually are.

**Prototypicality**

The effect sizes of the five prototypicality terms in the model were negligible with PRV values of either 0.000, 0.001, or 0.002. The largest effect was for diet product prototypicality (PRV = 0.002); this makes sense as this factor had a high effect in the WXYfm model (PRV = 0.123), thus being an ecologically valid predictor for food healthfulness. Regarding the category and prototypicality interaction, it seems that there are quite diverse effects depending on the food category. For categories such as 0.1 per cent milk, there are quite diverse effects depending on the food category and prototypicality interaction, it seems that more prototypical members of a category are perceived as more similar to each other (Loken, 2006) and thus different from other categories, whereas for the categories yogurt and cheese, the results seem in conflict with the theory.

How prototypicality can have such different effects depending on the food category is not clear at this point, and these results naturally call for further investigation, both to establish an accurate understanding of the effects of prototypicality and because this area could provide fruitful input for policy makers who wish to strengthen healthy food choices. It is particularly important to understand the relation between prototypicality and specific food categories because high levels of prototypicality could mislead consumers. With low levels of prototypicality, food categories such as yogurt and cheese are perceived as highly different (perceived healthfulness of about 4.3 vs. 3.2), whereas for high levels of prototypicality, these categories are perceived as similar (perceived healthfulness of about 4.0 vs. 3.8). High levels of prototypicality could therefore potentially lead consumers to believe that categories such as yogurt and cheese are equally healthful.

**Nutrition information**

The effects of nutrition values on consumer judgments were generally negligible, with effect sizes ranging from 0.000 to 0.004. Furthermore, there was almost no difference in the use of nutrition information between nutrition experts and nutrition illiterates. Although nutrition experts had a better understanding of nutrition and health, this did not translate into a higher use of nutrition information in judgments of food healthfulness.

**Labels and claims**

With regard to labels and claims, the study showed small effects of the organic label (PRV = 0.010) and of nutrition and health claims (PRV = 0.007) and no effect of the keyhole label (PRV = 0.000) on consumer judgments of food healthfulness. If organic labels and nutrition and health claims are used as cues for judging food healthfulness, this would be problematic as these cues had no effect on the WXYfm model (PRV = 0.000) and therefore do not reflect the healthfulness of foods.

**Implications for policy**

One of the main objectives behind the study was to clarify to what extent consumers utilize nutrition information in their judgments of food healthfulness. The results showed that nutrition information such as fat percentage, carbohydrates, sugar, or amount of calories had almost no effect on consumer judgments of food healthfulness, and this pattern was furthermore similar for nutrition experts and nutrition illiterates. A second important point is that the keyhole label, which is currently one of the most important policy instruments in promoting healthy food choice, did not have any effect on consumer judgments. These results are quite alarming as they seriously question the usefulness of both traditional back-of-pack nutrition labels and front-of-pack labels. What is even more alarming is that the nutrition experts and the nutrition illiterates were almost similar in that they do not base their healthfulness judgments on nutrition information. Furthermore, the high-knowledge group was only slightly more consistent in their evaluations and only slightly better at accurately judging the healthfulness of...
foods. This implies that educating the population to the highest level of nutrition and health knowledge will not change much in terms of judgmental accuracy. If education cannot change the way consumers make judgments, then it might be more meaningful to use consumer heuristics in a proactive way through, for instance, nudges or direct interventions such as differential pricing of healthful and unhealthful foods.

Limitations and future research
Several of the findings in this study call for further examination, particularly the findings on holistic package design and category prototypicality. With regard to holistic package design, further research is needed to establish whether consumers are actually misled by certain package designs or not. With regard to category prototypicality, further research is needed in order to establish a correct understanding of how prototypicality interacts with the food category in influencing healthfulness judgments. The study also raises doubt about the usefulness of nutrition labeling and the fact that increased nutrition and health knowledge does not translate into judgmental accuracy, which should be of special concern to policy makers.

Another issue that might deserve further attention is whether changing the perspective from individual food products to entire diets could affect consumer perceptions of food healthfulness. Cheese might, for instance, be considered an unhealthful food on its own but could easily be part of a healthful diet in combination with other foods. Does it make sense then to persuade consumers at the point of purchase to buy more healthful foods when the healthfulness of the food is actually decided at the point of consumption? Most policies seem to be directed at the level of food choices between similar alternatives; for example, the Danish keyhole label communicates that a particular food should be considered healthful and thus preferable within its category of foods. Such labeling schemes seem, however, to have a very limited effect on consumers, which makes one ponder whether consumers would not benefit from a different policy approach to food healthfulness.

Finally, an important limitation to the study is that the sample is restricted to Danish consumers. Previous research has demonstrated large intercultural differences in the understanding and conceptualization of foods (Rozin et al., 1999), and these, along with international variation in packaging design conventions, could play a role in consumer judgments of food healthfulness.

Appendix A: Examples of holistic package designs

<table>
<thead>
<tr>
<th>Natural</th>
<th>Contrasting</th>
<th>Delicate</th>
<th>Massive</th>
<th>Nondescript</th>
</tr>
</thead>
</table>

Appendix B: Nutrition and health claims from the product sample. Danish text in parentheses

(Let og lækker)
Light and delicious


(Ikke tilsat sukker. Indeholder sødestof og et naturligt indhold af sukker fra mælk og frugt) No added sugar. Contains sweetener and a natural content of sugar from milk and fruits.

(Hylld den gode smag uden at stikke dig. Økologisk og med lavt indhold af fedt) Celebrate the good taste without pricking yourself. Organic and with a low fat content.

(Hølde den gode smag uden at stikke dig. Økologisk og med lavt indhold af fedt) Keep the good taste without pricking yourself. Organic and with a low fat content.

(Balances your stomach)

(Holder maven i balance med 3 aktive probiotiske kulturer/Ikke tilsat sukker) Keeps the stomach in balance with 3 active probiotic cultures/no added sugar

(Omega 3 & 6. Fra raps- og solsikkeolie til stegning, bagning og madlavning) Omega 3 & 6. From rapeseed and sunflower oil for frying, baking and cooking

(Med vegetabilske frøoler med omega 3 & 6)
With vegetable seed oil with omega 3 & 6

Appendix C: Nutrition and health knowledge questionnaire

For the majority of Danish consumers, how many calories a day are recommended in a daily diet to maintain a healthy weight?

1. About 500.
2. About 1000.
3. **About 2000.**
4. About 3000.
5. About 3500.
6. Don’t know.
Response: 49.3 per cent correct.
Saturated fats are usually found in:
1. Vegetables and vegetable oils.
2. Animal products like meat and dairy.
3. Grain products such as bread and cereal.
4. None of the above.
5. Don’t know.

Response: 60.5 per cent correct.
Which kind of fat is higher in calories?
1. Saturated fats.
2. Polyunsaturated fats.
3. They are both the same.
4. None of the above.
5. Don’t know.

Response: 52.1 per cent correct.


Ministry of Food, Agriculture and Fisheries, & YouGov Zapera. 2010. Kendskab til nøglehulsmærket og kampagnen “nemt at vælge sundere”.
