Characterization of a food image stimulus set for the study of multi-attribute decision-making [version 1; referees: 2 approved]

Matthew Satterthwaite, Lesley K. Fellows

Department of Neurology and Neurosurgery, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada

Abstract

Everyday decisions are generally made between options that vary on multiple different attributes. These might vary from basic biological attributes (e.g., caloric density of a food) to higher-order attributes like healthiness or aesthetic appeal. There is a long tradition of studying the processes involved in explicitly multi-attribute decisions, with information presented in a table, for example. However, most naturalistic choices require attribute information to be identified from the stimulus during evaluation or value comparison. Well-characterized stimulus sets are needed to support behavioral and neuroscience research on this topic. Here we present a set of 200 food images suited to the study of multi-attribute value-based decision-making. The set includes food items likely to appeal to those accustomed to North American and European diets, varying widely on the subjective attributes of visual-aesthetic appeal ("beauty"), tastiness and healthiness, as rated by healthy young Canadian participants (N=30–67). The images have also been characterized on objective characteristics relevant to food decision-making, including caloric density, macronutrient content and visual salience. We provide all attribute data by image and show the extent to which attributes are correlated across the stimulus set. We hope this stimulus set will accelerate progress in the study of naturalistic, value-based decision-making.

Keywords

Value, choice, neuroeconomics, neuroaesthetics, eating behavior, obesity

Corresponding author: Matthew Satterthwaite (matthew.satterthwaite@mail.mcgill.ca)

Author roles: Satterthwaite M: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; Fellows LK: Conceptualization, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN 2016-06066) and the Canadian Institutes of Health Research (MOP-119291). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2018 Satterthwaite M and Fellows LK. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Satterthwaite M and Fellows LK. Characterization of a food image stimulus set for the study of multi-attribute decision-making [version 1; referees: 2 approved] MNI Open Research 2018, 2:4 (doi: 10.12688/mniopenres.12791.1)

Introduction

Value is a central concept in the field of economics, with rational decision-makers said to make choices that maximize subjective value. Neuroeconomics has provided evidence that the subjective value of options is represented in the brain, identifying neural signals in specific regions of the prefrontal cortex that are correlated with value ratings (Bartra et al., 2013; Clithero & Rangel, 2014), and predict choice (Kable & Glimcher, 2009; Montague & Berns, 2002; Wallis, 2007). Although there has been extensive work on global value representations in the brain, across a wide range of option types, little is known about how such representations are built up from the information provided by option attributes. Recent evidence indicates that attribute-values may be processed in parallel by distinct neural circuitry (Lim et al., 2013), and that different attributes may contribute to valuation over different time courses (Sullivan et al., 2015). Objective (e.g. calorie or macronutrient content) and subjective food-related attributes make distinguishable contributions to value, and these representations can be separated in the brain (Suzuki et al., 2017; Tang et al., 2014). Recently, prefrontal regions implicated in global value processing have been shown to be critical for allowing only some attributes to influence overall value or choice (Vaidya et al., 2017; Xia et al., 2015). Together, this work highlights the need for a more robust understanding of how stimulus attributes contribute to valuation.

Studies to date have generally concerned themselves with a limited number of attributes. For example, Lim et al. (2013) reported on how aesthetic and semantic attributes of logos on clothing influence consumer choice, while Xia et al. (2015) studied how competence and attractiveness of politicians contributed to voting choice. Recent work along these lines has begun to expand the number of attributes assessed, showing, for example, that the influence of 12 or more attributes can be discerned in the value ratings of visual artworks provided by healthy subjects (Vaidya et al., 2017).

Choices about food have been a recent focus of neuroeconomics research. Food choices are important to understand from evolutionary, health and economic perspectives. Food stimuli also offer a particularly rich range of potential attributes that might be expected to engage diverse brain processes. The subjective value of food has been shown to be influenced by several attributes, including taste and healthiness (Hare et al., 2009; Hare et al., 2011), aesthetic appeal (Van der Laan et al., 2012), caloric content (Tang et al., 2014), and constituent nutritive attributes such as protein, fat and carbohydrates (Suzuki et al., 2017). Additionally, basic perceptual attributes such as visual saliency (Mišosavljević et al., 2012) have also been shown to predict food choices in various circumstances. Thus, food choice likely draws on multiple attributes, ranging widely from signals relevant to biological drives, such as taste (Small et al., 2003) and nutrient content, to semantic attributes such as health, to more abstract considerations, such as aesthetic appeal.

Food stimulus sets have been developed by several groups (Blechert et al., 2014; Brodeur et al., 2010; Hare et al., 2009; Plassmann et al., 2007; ). Many neuroeconomics studies have used pre-packaged snack foods (i.e. Plassmann et al., 2007), providing limited range of health or aesthetic appeal. Others, (i.e. Hare et al., 2009) have varied the healthiness of items, but presented them on white or black backgrounds, limiting the aesthetic range. Here, we present a food image stimulus set developed to study multi-attribute decision-making across a wide range of attributes. The aim was to provide diverse, appetitive food options varying widely in subjective ratings of taste, health, and aesthetic beauty, as well as objective indicators of caloric density, constituent macronutrients and saliency.

Methods

Selection of stimuli

The stimulus set consists of 200 colour photographs of foods commonly eaten in North America, (referred to here as the “Full set”) covering a wide range of items from savory (e.g., pasta, steak), sweet (e.g., cake, ice cream), processed (e.g., pretzels, chips), to natural (e.g., strawberries, salad) food items. Images could depict single or multiple items, depending on typical serving sizes of the particular food. Images did not have any brand marker or logo; the food is shown free of any packaging. A total of 180 images were initially selected, with an additional 20 images added to increase stimulus set size and attempt to reduce correlations between attributes (discussed further in the “Subjective attribute rating” section).

Images were drawn from a variety of online sources. 105 were stock images free of copyright (referred to here as the “Open set”), which are provided in the “Open stimulus set” folder, and are available on OSF (Satterthwaite, 2018). We expected that stock images would not provide sufficient variability in aesthetic value; thus an additional 95 images were drawn from other sources that did not explicitly waive copyright. Our understanding is that such images can be used for research purposes. However, to respect copyright, these images are not included with this article. They can be acquired for research use by writing to the first author. The copyright status of each image is provided in the “Image attributes” spreadsheet.

Image customization

All images were cropped using Adobe Photoshop CS6 to 250 by 250 pixels.

Participants

The stimulus set was rated by a convenience sample of 67 healthy adult participants (28 men, 39 women; mean age (SD): 21.3 (2.6) years; range, 18–29 years), consisting primarily of McGill University undergraduate students recruited by advertisement, as part of an in-lab experimental session conducted between June 2017 and February 2018. Exclusion criteria were self-report of any neurological or psychiatric disease, eating disorder, or current use of psychoactive drugs. All subjects had normal or corrected-to-normal vision. Height and weight were measured for 47 of the participants to calculate body mass index (BMI). Using the BMI classification of the World Health Organization, 35 participants were classified as being in the normal range of BMI, 2 were moderately to mildly thin, 8 were overweight and 2 were obese. We also recorded
whether participants had any dietary restrictions (e.g. vegetarianism, food allergies). In order to examine test-retest reliability, 11 participants completed the ratings again in a second session approximately 90 days later. Demographic information is summarized in the “Participant Information” spreadsheet. The study was approved by the McGill University Health Centre Research Ethics Board (NEUPSY-# 2017-26190), and all participants provided written informed consent.

Subjective attribute rating
Participants rated a given attribute for each image in the stimulus set, then viewed the set again to rate the second, and then the third attribute. They were asked to indicate how “tasty”, “healthy” and “aesthetically pleasing” each food item was by clicking on a scale from 1 to 7 with 1 indicating “not at all” and 7 indicating “very much” (Figure 1). Participants were given as much time as they needed to choose a rating, and were instructed to use the whole scale. The order of images was randomized within each task, and the order of the rating tasks was randomized across participants. The first 180 images (image number 1 to 180) were rated by all 67 participants. Interim analysis led to the addition of 20 images to the stimulus set, which were selected in an attempt to reduce correlations across attributes. The additional images (#181-200) were rated by 30/67 participants, who completed the experiment after the addition of the novel stimuli. Rating tasks were implemented in E-Prime 2.1 (Psychology Software Tools, Inc., Pittsburgh, PA, USA) with stimuli presented on a ViewSonic VX2370Smh-LED screen with a refresh rate of 60 Hz. Participants were seated in a quiet testing room, approximately 60 cm from the computer screen.

Characterization of calorie and nutrient content
Caloric density and macronutrient content for each item were estimated using Conde Nast’s “nutritiondata.self.com” website. Caloric density is reported in units of calories per 100 g, while protein, fat and carbohydrate content are reported in grams per 100 g of food.

Saliency
Saliency scores for each image were calculated using the default parameters of the Saliency Toolbox (Walther & Koch, 2006).

This Matlab toolbox evaluates the saliency of images based on equal weightings of intensity, color and orientation.

Data analysis
All statistical analyses (ANOVA with post hoc Bonferroni correction or Pearson correlation; described in the Results section) were performed using IBM SPSS Statistics 23. Unless otherwise indicated, the results refer to the full set of 200 images.

Results
Subjective attributes
The mean attribute ratings across the stimulus set are shown in Table 1 and the distributions of mean image ratings for each attribute are shown in Figure 2. Individual image ratings from each participant are provided in the “Ratings_all” spreadsheet, while the mean and standard deviation of attribute rating per image are provided in the “Image attributes” spreadsheet. As intended, the images in the stimulus set varied widely on each attribute. Mean ratings of individual stimuli ranged from 1.9 to 5.8 for taste, 1.2 to 6.9 for health, and 1.8 to 5.9 for aesthetic beauty (Figure 2). The standard deviation in ratings for individual images ranged from 1.1 to 2.2 for taste, 0.3 to 1.6 for health, and 1.1 to 2.0 for aesthetic beauty. A one way repeated measures ANOVA for the effect of attribute on the standard deviation of individual stimuli ratings indicated that there was a significant effect of attribute on rating standard deviations (F(2,398) = 546.30, p < 0.001). Post hoc pairwise comparisons indicated that stimuli standard deviations for health ratings were lower than for both taste and beauty ratings (Bonferroni tests, p < 0.001) and that standard deviations for beauty ratings were lower than for taste ratings (Bonferroni test, p < 0.001). Thus, the agreement of ratings for a given image, across subjects was highest for health ratings and lowest for taste ratings.

Objective attributes
The ranges of objective attributes across the stimulus set are shown in Table 1, with distributions of caloric density, macronutrient content and saliency shown in Figure 3. The characteristics of each image are provided in the “Image attributes” spreadsheet. As intended, the stimulus set also varied widely on caloric and nutritive density. Individual characteristics ranged...
Table 1. Subjective attribute ratings and objective characteristics across the full stimulus set and the open stimulus set. Data are presented as the mean ± standard deviation.

<table>
<thead>
<tr>
<th>Health</th>
<th>Beauty</th>
<th>Taste</th>
<th>Caloric density (calories/100g)</th>
<th>Protein (g/100g)</th>
<th>Fat (g/100g)</th>
<th>Carbohydrates (g/100g)</th>
<th>Salience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full set (n=200)</td>
<td>3.9 (1.7)</td>
<td>4.2 (1.0)</td>
<td>4.4 (0.8)</td>
<td>244 (152)</td>
<td>9 (8)</td>
<td>11 (11)</td>
<td>28 (24)</td>
</tr>
<tr>
<td>Available here (n=105)</td>
<td>4.2 (1.7)</td>
<td>4.3 (0.9)</td>
<td>4.4 (0.7)</td>
<td>212 (139)</td>
<td>9 (8)</td>
<td>10 (10)</td>
<td>23 (22)</td>
</tr>
</tbody>
</table>

Figure 2. Distribution of mean subjective attribute ratings for health, beauty and taste for both the full set and the open set of images.

Figure 3. Distribution of protein, fat and carbohydrate content, as well as caloric density and saliency score distributions for the full set and open set of images.
from 16 to 654 calories/100 g, 0.0 to 38.3 g/100g of protein, 0.0 to 65.2 g/100g of fat, 0.0 to 98.9 g/100g of carbohydrate and 8.6 to 41.8 for image salience (see Figure 3).

Correlation between attributes
Pearson correlations between the mean subjective ratings (taste, health, aesthetic beauty) and objective attributes (caloric density, protein, fat, carbohydrate, salience) of all 200 images are shown in Table 2.

Test-retest reliability of subjective attribute ratings
The reliability of subjective attribute ratings across two test sessions was assessed in 11 participants tested at least 90 days apart (mean, 118 days; range, 90–267 days). We looked at the correlation of within-subject attribute ratings for each image on each attribute. Ratings for all three attributes were moderately to strongly correlated across test sessions (aesthetic beauty, r(1980)=0.62, p<0.01; health: r(1980)=0.86, p<0.01; taste: r(1980)=0.75, p<0.01).

Demographic effects
We next tested whether mean ratings across attributes differed by gender, age or BMI. A 1 within (attribute task) 1 between (gender) repeated-measures ANOVA determined that there was no effect of gender on ratings for any of the subjective attributes (p >0.05). Pearson correlations indicated that there were no correlations between age and any of the attribute ratings (p values >0.05), BMI and health ratings (p = 0.25) or BMI and taste ratings (p = 0.93). There was a significant negative correlation between BMI and beauty ratings (r(47) = -0.34, p = 0.021). This was mainly driven by the participants with BMI in the obese range: when these two participants were removed from the analysis the correlation was no longer significant (r(45) = -0.15, p = 0.33).

The majority (54/67) of our participants did not report any dietary restrictions. The most common restriction was vegetarianism (n=5), but this small sample size did not allow us to detect any systematic variation in how these individuals rated the stimuli.

Discussion
We present a food image stimulus set characterized on three subjective attributes and five objective attributes potentially relevant to value-based decision-making. As intended, these images varied widely on subjective attributes of aesthetic beauty, healthiness and tastiness, as rated by young, healthy participants. The images also varied on objective attributes of caloric density, nutrient (protein, fat, carbohydrate) content and visual saliency. We found that subjective attribute ratings across participants were most consistent for health, followed by beauty and then taste. We also found moderate to high test-retest reliability for these ratings over a 90-day delay.

Studies investigating the influence of attributes on decision-making have generally collected attribute ratings from participants as part of the experiment (such as in studies by Hare et al., 2009 and Sullivan et al., 2015). This can be time-consuming, and may potentially bias individuals to consider attribute information that they would not have otherwise considered if the value rating or decision-making phase of the experiment follows (for example, in designs such as that by Hare et al. (2009) where choice pairs were created based on specific combinations of attribute ratings). Additionally, requiring participants to rate the same image multiple times on different attributes reduces the novelty of the stimulus set, which can be particularly important in paradigms where individuals are required to make quick decisions based on limited exposure to the stimuli (as in the study by Milosavljevic et al., 2012). The mean ratings we provide here could be used to avoid these issues; this dataset allows for selection of the most consistently rated images. However, unsurprisingly, variability is evident in the subjective ratings across participants, particularly for taste and beauty ratings, and this needs to be considered in experimental design.

In selecting the images, we aimed to minimize correlations across attributes while still providing a range of familiar foods. Based on prior work (Scheibehenne et al., 2007) we expected at least moderate correlation of some attributes, as this is a feature of naturalistic food stimuli. Work on everyday decision-making argues that such correlations may serve to simplify and speed decisions, and thus could be harnessed for work on heuristic food choice (Gigerenzer & Goldstein, 1996). However, if the aim is to disentangle attribute contributions to value (e.g. Hare et al., 2011; Lim et al., 2013; Suzuki et al., 2017), these

Table 2. Correlation coefficients for all subjective and objective attributes.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Beauty</th>
<th>Health</th>
<th>Taste</th>
<th>Caloric density</th>
<th>Protein content</th>
<th>Fat content</th>
<th>Carbohydrate content</th>
<th>Saliency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beauty</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health</td>
<td>0.349**</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taste</td>
<td>0.544**</td>
<td>0.006</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caloric density</td>
<td>-0.305**</td>
<td>-0.640**</td>
<td>-0.122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein content</td>
<td>-0.061</td>
<td>-0.076</td>
<td>0.166*</td>
<td>0.308**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat content</td>
<td>-0.148*</td>
<td>-0.376**</td>
<td>0.030</td>
<td>0.814**</td>
<td>0.0393**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbohydrate content</td>
<td>-0.307**</td>
<td>-0.569**</td>
<td>-0.285**</td>
<td>0.630**</td>
<td>-0.285**</td>
<td>0.130</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saliency</td>
<td>0.164*</td>
<td>0.115</td>
<td>0.117</td>
<td>0.010</td>
<td>0.108</td>
<td>0.096</td>
<td>-0.129</td>
<td></td>
</tr>
</tbody>
</table>

**P < 0.01, *P < 0.05.
correlations should be kept in mind when selecting stimuli from this set.

Here we provide ratings made by a convenience sample of healthy, young, well-educated participants living in Canada. We saw no consistent effect of age, gender or dietary restrictions on attribute ratings. Although most had a BMI within the normal range, we observed that BMI was negatively correlated with beauty ratings, primarily driven by the participants with the highest BMI. Obesity has previously been shown to affect how choices are influenced by perceived healthiness of foods (Lim et al., 2018), but to our knowledge this is the first evidence that BMI may also relate to the perceived aesthetic beauty of food options. Future work might usefully address how this relates to food valuation and choice. The generalizability of the subjective attribute ratings to older or clinical samples, including people with more severe obesity, or to people accustomed to different foods in their usual diet also remains to be determined.

Overall, the stimulus set that we present here varies widely on a number of subjective and objective attributes. The open set (available for download here), provides adequate variability on most attributes and should suffice for the majority of decision-making paradigms. For paradigms requiring larger stimulus sets or more stringent control of presented attributes, the full set (available upon request) should provide a useful starting point.

Data availability

Copyrighted images used in this study can be obtained for research use by emailing the corresponding author (matthew.satterthwaite@mail.mcgill.ca).

Raw data for the study and the uncopyrighted imaged are available on OSF: https://doi.org/10.17605/OSF.IO/5GRBE (Satterthwaite, 2018).

Competing interests

No competing interests were disclosed.

Grant information

This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2016-06066) and the Canadian Institutes of Health Research (MOP-119291).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.


Open Peer Review

Current Referee Status: ✔ ✔

Kiyohito Iigaya, Logan Cross
California Institute of Technology, Pasadena, CA, USA

In this methods article, the authors described a new dataset about attributes ratings on various food images. The attribute sets include subjective ratings (health, beauty, taste) and objective characteristics (caloric density, protein, fat, carbohydrates, salience). The description of the dataset is clear, and the dataset is potentially very useful for understanding value-based decision-making. We have a couple of comments:

- The dataset includes reaction time, which we think potentially very interesting to look at. So we suggest the authors to mention reaction time in the article, and perhaps include histograms or a table, as they did so for ratings.

- Labels/numbers in Figures 2 and 3 are too small to read.

Is the rationale for developing the new method (or application) clearly explained?
Yes

Is the description of the method technically sound?
Yes

Are sufficient details provided to allow replication of the method development and its use by others?
Partly

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
Satterthwaite and Fellows provide a set of food images along with information about multiple attributes of the food items (e.g., caloric density, macronutrient content and visual salience). I believe the data set is to be of wide interest to neuroscientists, psychologists and other researchers interested in human decision-making. I have only one minor comment on the manuscript. As the authors mentioned, the data set is relevant only for those who are accustomed to North American and European diets. It would be great if they could briefly discuss the limitation and possible cultural-differences.

**Is the rationale for developing the new method (or application) clearly explained?**
Yes

**Is the description of the method technically sound?**
Yes

**Are sufficient details provided to allow replication of the method development and its use by others?**
Yes

**If any results are presented, are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions about the method and its performance adequately supported by the findings presented in the article?**
Yes

**Competing Interests:** No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.