Consumer expectations for vegetables with typical and atypical colors

The case of carrots

Schifferstein, Hendrik N.J.; Wehrle, Theresa; Carbon, Claus-Christian

DOI

Publication date
2019

Document Version
Final published version

Published in
Food Quality and Preference

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.
Consumer expectations for vegetables with typical and atypical colors: The case of carrots

Hendrik N.J. Schifferstein\textsuperscript{a,c}, Theresa Wehrle\textsuperscript{b,c}, Claus-Christian Carbon\textsuperscript{b,c}

\textsuperscript{a} Department of Industrial Design, Delft University of Technology, the Netherlands
\textsuperscript{b} Department of General Psychology and Methodology, University of Bamberg, Germany
\textsuperscript{c} Research Group EPÆG (Ergonomics, Psychological Aesthetics, Gestalt), Bamberg, Germany

Abstract

The variety of fruits and vegetables in today's supermarkets is enormous. We investigated how color may lead consumers to anticipate differences in product properties. Forty volunteers rated the expected properties for carrots with different colors presented in pictures, together with their familiarity, purchase intention, and intended preparation method on 7-point scales. We found most positive expectations for the typical and most familiar kinds of orange carrots. Lower saturation of orange was associated with lower attractiveness and freshness, whereas higher orange saturation was evaluated as more artificial. Brown spots on carrots were associated with disease and such carrots were regarded less healthy. Carrots in atypical colors were rated as less familiar, attractive and healthy than orange ones. In comparison with the orange carrots, red carrots were expected to taste sourer and spicier, purple and yellow carrots were rated less nutritious and more artificial, with purple carrots expected to taste more bitter and yellow ones more sour. White and white-green carrots rated lower on sweet, and higher on sour, bitter, and spicy. These carrots were considered less ripe and less nutritious than orange ones. These results indicate that color hue and saturation have substantial impact on consumers' expectations about sensory and functional properties, including freshness and nutritional value. Some of these expectations may be derived from associations to other vegetables, as reflected by high ratings for spiciness (red pepper) and taste intensity (turnip, radish). However, low attractiveness ratings also suggest that consumers may be reluctant to try unfamiliar variants, at least at first glance. Although atypical colors produce culinary opportunities, commercial success may be limited until consumers integrate them in their everyday habits.

1. Introduction

Have you ever wondered what a purple carrot tastes like? The variety of fruits and vegetables on display in today's supermarkets is enormous. Diversity within each species has increased considerably too, with differences in size, shape, color, flavor, production and trading method. Because of the increased variation that is available to consumers, many of the product varieties offered are likely to be unfamiliar to the naïve consumer, leading to uncertainty about what qualities and taste experiences to expect.

Even though a specific color variety may be unfamiliar to the majority of current consumers, the variety may be familiar to experts who breed and grow these vegetables. Banga (1963) writes about carrots: “Purple varieties were known in Europe until the beginning of the 20th century. They are still popular in Egypt, Asia Minor, India, Japan, China, and other Asiatic countries today […]. Yellow varieties used to be very popular in Europe as a winter food, but they have lost most of this importance now. Yellow varieties are also known in Eastern countries. White varieties were sometimes used for human food in European countries in the 18th and 19th centuries, but more for the feeding of cattle. In Eastern countries some varieties are white. Orange varieties now predominate in the Western world. In the East they are found together with purple, yellow and white varieties” (p. 357). Although this account was written more than fifty years ago, it clearly shows that the familiarity with different color varieties has changed over time and may vary between geographical regions.

In the present study, we investigate how variation in color may lead consumers to anticipate differences in product properties. Obtaining insight into such variations in expected properties is important for vegetable producers and retailers, because these expectations are likely to influence consumers’ judgments of appropriateness for the use in particular dishes, the preparation methods they employ, their perception during actual consumption, and last but not least, the consumers’ purchase probability.
We will now discuss how different colors may affect the expected sensory properties of fresh vegetables. We specifically focus here on the literature investigating color effects for fresh agricultural products, because perception and evaluation mechanisms may be quite different for highly processed foods.

2. The effect of color in interactions with vegetables

Consumers make use of the color of vegetables as a cue to identify the product, to assess the safety, quality and ripeness of the product, and to make inferences about its sensory properties. For many vegetables, people are familiar with prototypical colors (e.g. orange for carrots, red for tomatoes, green for cucumber) and these colors make it easy to identify the product class and thus support consumers while shopping. But what happens if the colors of the products they see deviate from what they are used to?

Relatively subtle color differences (e.g., several shades of red for tomatoes) may already have a substantial impact on expected food preference and food quality evaluations (e.g., Bruhn, 1995; Crisosto, Crisosto, & Metheney, 2003), because these color changes may be interpreted as meaningful information. For instance, they may be used as indicators of product freshness, ripeness, or nutrient content (e.g., De Groote & Kimenju, 2008; Lekrisompong, Whitson, Truong, & Drake, 2012), with implications for expected sensory properties, such as sweetness, sourness, bitterness, hardness, and so on. Lee, Lee, Lee, and Song (2013) conducted an experiment investigating the effect of color manipulations for pictures of multiple natural foods. They generally found a consumer preference for highly chromatic food colors, a tendency that was largest for foods with red and green hues. As consumers intuitively relate the intensity of colors of foods to the intensity of flavors (Calvo, Salvador, & Fiszman, 2001; Christensen, 1983), this suggests that they may have a preference for foods that are rich in flavor.

Larger color deviations are likely to have an even bigger impact on consumer expectations. For instance, a tomato that is white or green instead of red, may lead consumers to conclude that the product is unripe (except for some geographical idiosyncratic differences, where green tomatoes are part of the established cuisine). Hence, in the unprocessed state consumers probably expect to taste a hard, sour or bitter product. Commonly, consumers will follow the general rule that the redder the tomato is, and the darker and more saturated its color becomes, the sweeter and softer the product will be. Additionally, when the color becomes browner or the vegetable contains brown spots, the consumer is likely to conclude that the product is decaying, and will expect the product to be softer, bitterer, and possibly smell bad. Lee et al. (2013) analyzed pictures of several decaying natural food products and found that the colors of most products generally shifted in the direction of the less chromatic colors, with bananas also becoming less greenish and lettuce, cucumber, and red apples all three becoming more yellowish over time. Analogously, consumers would likely take beef meat to have a red color, whereas a grey, brown or purple color may indicate a lack of freshness and a potentially easily perishable product, which lowers consumer purchase intentions (Acebrón & Dopico, 2000; Carpenter, Cornforth, & Whittier, 2001). de Hooge et al. (2017) investigated the relevance of color deviations by presenting a picture of an apple that had a suboptimal appearance due to a brown spot. The associations that this apple generated included: to be discarded, bad tasting, unattractive to eat, unsafe to eat and to be used in cooking. de Hooge et al. found that such apples are likely to obtain lower liking ratings than unaffected apples and will be bought by only few consumers.

In cases where color differences are large and a color is atypical for the product in the sense that it cannot be related to the usual ripening or decay processes (e.g., a blue or purple color for a tomato), the consumer is likely to conclude that this product is not a regular product. Probably, being unfamiliar with the product will already lead to lower acceptance ratings (e.g., Tuorila, Meiselman, Bell, Cardello, & Johnson, 1994). Lekrisompong et al. (2012) determined the acceptability of orange, yellow, and purple cultivars of sweet potatoes in consumer taste tests with or without blindfold. As expected, the study showed that the most common type of sweet potatoes in the US research population (orange) received higher liking scores compared with yellow or purple cultivars. However, in this study the effects of using the blindfolds were minimal, suggesting that these liking differences were not due to the color differences per se. Paakkilä, Sandell, and Hopia (2016) prepared two kinds of potato salads, one with yellow- and one with blue-fleshed potatoes. The taste of both salads was the same, even though the color deviated. However, the majority of participants (64.7%) preferred the yellow-fleshed and just 28.1% the blue one. Also, the yellow one received higher ratings at every attribute except bitterness. The negative effects of unfamiliarity, however, might be particularly present for adults and not be found in groups of children: Some experiments have revealed that for children differences in color had no significant effects on acceptance and that atypically colored samples were actually preferred significantly more often by appearance (Poelman & Delahanty, 2011), an effect that could be attributed to less stable and established representations of norms.

If vegetables have unfamiliar colors, consumers may be unsure about the nature of the product and they may confuse or associate it with other products sharing certain properties. For instance, if they see a white vegetable in the shape of a carrot, they may be reminded of a turnip and they may guess that the qualities of the white carrot may approach those of a turnip. Probably, the type of associations with other products will largely determine the flavor expectations they have, because people tend to have only few general color-taste associations (Koch & Koch, 2003) and many associations tend to be highly product specific (Mage, 1974). For instance, Shermer and Levitan (2014) found that the darkness of the red color of salsa was directly related to the expected spiciness of the sauce, as linked to associations with red peppers. For other types of products, different associations may be more appropriate with the same color. For instance, some studies have suggested that a red color may enhance sweetness ratings in beverages (Johnson & Clydesdale, 1982; Johnson, Dzendzolet, & Clydesdale, 1983), probably due to associations with red fruits like strawberries and cherries.

2.1. The present study

For the present experiment, we aimed to test color deviations of carrots in an ecologically valid context. To follow this aim, we focused on naturalistic depictions of real-life products instead of typical experimental, highly controlled variations of particular aspects of the material. This is a fundamentally different approach to experimental procedures where most effort is put into the independent crossing of several variables, such as the properties of a product. We have deliberately decided against such a standard approach in order to capture assessments that are grounded in and are relevant to real-life consumer behavior. Nonetheless, we standardized the way in which the materials were presented, so that stimuli were identical for all participants.

To test a very wide selection of carrots, we used two sets of pictures of carrots covering natural representations of various colors of carrots that are available in supermarkets nowadays. In the first set of carrots, the base color of all carrots was orange, but the nuances of the color deviated. However, the majority of participants (64.7%) preferred the yellow-fleshed and just 28.1% the blue one. Also, the yellow one received higher ratings at every attribute except bitterness. The negative effects of unfamiliarity, however, might be particularly present for adults and not be found in groups of children: Some experiments have revealed that for children differences in color had no significant effects on acceptance and that atypically colored samples were actually preferred significantly more often by appearance (Poelman & Delahanty, 2011), an effect that could be attributed to less stable and established representations of norms.
saturation values and more browning to decrease in quality and to decomposition, and thus to lower acceptance ratings (de Hooge et al., 2017; Lee et al., 2013), and possibly higher ratings for bitterness and lower ratings for product firmness, due to loss of structure.

In the second set of carrots, orange carrots were compared to carrots with atypical colors. Because atypical colors for food products generally result in lower acceptance ratings in groups of adults, we predict that carrots with atypical colors will receive lower attractiveness ratings than orange colored carrots. In addition, in the case of white and green carrots, the colors may suggest that the carrots are unripe. Hence, we expect them to have higher ratings for attributes such as crunchiness, bitterness, and sourness and lower ratings for ripeness and nutritional value. Furthermore, among the atypical hues we expect that consumers may have associations with various types of other vegetables that will bias their expectations. For instance, red carrots may evoke associations to red peppers (Shermer & Levitan, 2014) and thus we expect them to be evaluated as more spicy. White carrots may trigger associations to white radish and horseradish, and thus we expect them to be rated higher on spiciness as well.

3. Method

3.1. Participants

Forty volunteers, 23 females and 17 males, of many different regional nationalities (50% European, 45% Asian, 5% South American) participated in the experiment. Ages varied between 19 and 28 years ($M_{\text{male}} = 24.2$ years, $SD = 2.1$; $M_{\text{female}} = 24.5$ years, $SD = 2.7$). The majority were undergraduate (75%) and graduate (13%) students of Delft University of Technology. They were compensated with EUR 10 for their participation.

Consuming carrots is very common in the Netherlands. Carrots were among the ten most eaten vegetables during the main course in 2011. They are mainly bought fresh, either uncut (55%) or precut (21%), and sometimes precooked in a jar (12%), a can (4%), or frozen (4%). The most common preparation method is cooking (51%), but a significant amount is eaten raw, either on its own (6%) or in a salad (12%). Other preparation methods include stir-frying (10%), baking (4%) and steaming (4%) (Borgdorff, 2012).

In the current sample, all participants reported that they ate vegetables on a regular basis: 80% often and 20% sometimes. Most of them liked eating carrots, with only 20% giving a liking rating of 4 or lower on a 7-point scale ($M = 5.3$, $SD = 1.45$). On average, participants estimated that they consumed 371 g cooked, 290 g raw, 103 g mashed, 15 g baked carrots, and 198 g carrots in soup per month.

3.2. Stimuli

Pictures of differently colored carrots were obtained through the Internet. We looked for images with multiple carrots that were similar in shape and size, and that were lying or hanging together, preferably with a small bit of green leaves at the top. In addition, we searched for images that were large in size and of good photographic quality. If necessary, we adapted these pictures to match our requirements.

We chose to use pictures from the Internet rather than taking our own pictures of carrots, because this gave us access to a large number of images with carrots varying in shapes, lengths, types and colors. This enabled us to select images with carrots that differed in color, but had similar shapes. Taking our own pictures of many different carrots would be unlikely to provide us with all the varieties we were looking for. Obtaining carrots in many different colors with similar shapes from local supermarkets would be hard, given that local stores have limited selections and that many varieties will not be available year-round. Hence, searching the Internet was likely to provide us with more variety, as it would reflect world-wide availability from many areas and seasons.

Furthermore, as we wanted to investigate only existing color variations of carrots, we did not chose to use a graphical manipulation or 3D model approach by artificially coloring existing pictures or models of carrots. First of all, we were unsure whether we would be able to recreate the exact colors that were available on the market. Second, our initial attempts showed that pictures derived in this way looked quite artificial, which would likely distract participants from giving veridical evaluations. Most importantly, we expected that an experiment where multiple pictures would be presented showing exactly the same shape of carrot, but in many different colors would raise suspicion among our participants about the aim of the study. All in all, we decided to implement an ecologically valid instead of a perfectly crossed experimental setting to maximize the validity of the assessments. The resulting pictures yielded exemplars that differed in photographic perspective, lighting conditions, and the backgrounds on which carrots were displayed. The obtained pictures, however; also looked very natural, like pictures from a purchase catalogue on carrots. Although this procedure resulted in less controlled experimental conditions, we gained very authentic material to which participants could devote their full concentration and motivation for valid assessment.

Four pictures contained orange carrots: Three of these pictures showed carrots in uniform orange color with low, intermediate, and high degree of saturation and the fourth picture showed orange carrots with a number of brown spots (Fig. 1). Five additional pictures showed carrots with atypical colors (Fig. 2). On four of these pictures, the carrots were uniformly colored in a single hue: purple, red, white or yellow. The fifth picture showed carrots that were mainly white, with a patch of green at the top. We used the latter picture, because we expected the green patch to suggest that the carrots were unripe.

In order to quantify the color properties of the carrots in the pictures, we plotted 50 randomly selected local HSV (hue, saturation value) color patches of each carrot on a color wheel in the program R. In order to obtain the 50 dots, we asked a person who was naive with respect to the purpose of the study to select 50 dots that covered the entire spectrum of the areas of the carrots. We preferred this procedure rather than using an automatic cropping algorithm, because the corners and edges would be highly ambiguous in terms of correspondence to figure or ground, and the shadows there would interfere with the measurement. The color wheels show hues from $-180^\circ$ to $+180^\circ$, with $0^\circ$ representing red at $12$ o’clock. Saturation information can be retrieved from eccentricity: the more saturated the color, the more eccentric the representation on the wheel, with lowly saturated colors represented by centrally oriented dots. For instance, the white carrot variant shows dots that are strongly clustered near the center of the color wheel. Value (lightness) information is encoded as the size of the dots: the lighter the color of the measured area, the larger and lighter the dot.

The analysis of the color values in Fig. 3 and Table 1 confirms that the pictures of the orange carrots indeed vary in the saturation of the orange color, with the value of the high saturation carrots (0.803) exceeding the values for medium (0.741) and low (0.673) saturation carrots. Whereas the numbers for value are very close, the hue seems to differ a bit, with the hue value for the medium saturation carrots exceeding the values for the low and high carrots. This suggests that the medium carrots on average may be a bit more yellowish than the high and low saturation carrots.

3.3. Procedure

The study was conducted in a laboratory with no windows and only artificial lighting. Given that the study was conducted with people from various nationalities, the entire study was conducted in English. Sessions were conducted separately for each respondent. After entering the room, participants filled in a consent form and performed a test of visual acuity and color vision. After passing these tests, participants provided demographic information and filled out a questionnaire...
Fig. 1. Pictures with orange carrots used as stimuli with low (334), medium (167), and high (428) saturation of the orange color, and with brown spots (526). The authors thank Pat Herman (167; see freimages.com), Kander (334; copyright according to wikimedia.org), Hild Samen Gmbh (428; www.hildsamen.de), and Rasbak (526; copyright according to wikimedia.org) for permission to reproduce the pictures. Links to all pictures can be found online through https://nl.pinterest.com/rickschierrerste/carrots/. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 2. Pictures of carrots in different colors: orange (167), red (193), yellow (270), purple (879), white-green (914), and white (603). The authors thank Pat Herman (orange (167) freimages.com), Osborne Seed Company (yellow (270) from osborneseed.com), Nunhems BV (purple (879), and crème delight (603) from nunhems.com), and Vreeken’s zaden (white Belgian (914) from vreeken.nl) for permission to reproduce these pictures. Unfortunately, despite extended efforts, we were unable to trace the copyright owner of the atomic red carrots (193) picture. Links to online sources can be retrieved through https://nl.pinterest.com/rickschierrerste/carrots/. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
regarding their general eating habits and their consumption of carrots. Subsequently, they were seated at a distance of 50 cm from a computer monitor (EIZO Color Edge CG 277, 27-inch diagonal) that was color-calibrated with Color Navigator 6 before every experimental session in order to have identical and also true color settings at each session. The nine experimental images were presented at the screen in a random order that differed between participants. Each image was rated using a paper and pencil questionnaire.

In order to develop a questionnaire that covered the most important topics that consumers considered when they evaluated carrots, we performed a pre-study for which we recruited ten volunteers (6 females and 4 males of different nationalities, aged between 24 and 30, $M_{\text{male}} = M_{\text{female}} = 26.7$ years) from the same participant pool. None of these participants took part in the main study. We conducted two focus group discussions with five participants each. We presented them with twelve pictures of carrots, similar to the ones used in the main study. We asked them to indicate which properties they had perceived while eating the different types of carrots in the past, and what properties they associated with the different carrots, either from personal experience or based on their expectations. We also asked them whether they had particular associations (symbols, people, sceneries, situations) to the various carrots. On the basis of the notes and audio recordings of the discussions in the two groups, we generated a list of attributes that were mentioned and we clustered these. Twenty-three questions were

<table>
<thead>
<tr>
<th>Hue</th>
<th>Satur.</th>
<th>Value</th>
<th>Hue</th>
<th>Satur.</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>High saturation</td>
<td>17.01</td>
<td>0.631</td>
<td>0.803</td>
<td>0.010</td>
<td>0.895</td>
</tr>
<tr>
<td>Medium saturation</td>
<td>23.64</td>
<td>0.669</td>
<td>0.741</td>
<td>0.020</td>
<td>0.863</td>
</tr>
<tr>
<td>Low saturation</td>
<td>15.23</td>
<td>0.670</td>
<td>0.673</td>
<td>0.023</td>
<td>0.849</td>
</tr>
<tr>
<td>Brown spots</td>
<td>19.09</td>
<td>1.187</td>
<td>0.675</td>
<td>0.028</td>
<td>0.732</td>
</tr>
<tr>
<td>Red</td>
<td>–1.72</td>
<td>2.012</td>
<td>0.778</td>
<td>0.026</td>
<td>0.716</td>
</tr>
<tr>
<td>Yellow</td>
<td>48.37</td>
<td>0.520</td>
<td>0.711</td>
<td>0.020</td>
<td>0.801</td>
</tr>
<tr>
<td>Purple</td>
<td>–35.23</td>
<td>3.399</td>
<td>0.507</td>
<td>0.028</td>
<td>0.753</td>
</tr>
<tr>
<td>White</td>
<td>51.09</td>
<td>5.489</td>
<td>0.204</td>
<td>0.025</td>
<td>0.762</td>
</tr>
<tr>
<td>White-green</td>
<td>58.44</td>
<td>6.311</td>
<td>0.205</td>
<td>0.033</td>
<td>0.771</td>
</tr>
</tbody>
</table>

Fig. 3. Representation of the carrots in the nine stimuli on color wheels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
developed for the main study, based on their judged relevance for the manipulated variables (color saturation, typicality, carrot freshness and ripeness).

In the main experiment attributes were assessed on 7-point, unnumbered Likert scales. The participants answered the question “I find that this carrot looks...” for seven unipolar attributes (familiar, attractive, healthy, crunchy, sweet, bitter, sour) on scales from 1 (not at all) to 7 (very) and for eight bipolar attributes (weak – strong in taste, old – fresh, dry – juicy, mild – spicy, stiff – flexible, nutrient-poor – nutrient-rich, natural – artificial, unripe – ripe) moderated with very on either side of the scale. They also indicated to what extent the carrots they had tasted so far were similar to the carrot in the picture (very dissimilar – very similar) and how high the probability was that they would buy the carrot (very improbable – very probable). Furthermore, they indicated the probability that they would prepare and eat the carrot in different modes (raw, cooked, mashed, in a soup, baked) on a 7-point scale from very improbable to very probable. Finally, they answered an open question ‘What does this carrot remind you of?’

Time of image exposure was unlimited and ended by the participant’s keypress. In total, every participant finished nine runs in order to rate all different images. A typical session lasted approximately 45 min.

3.4. Data analysis

In a doubly multivariate analysis of variance, we first tested for the impact of carrot type (9 levels) as within-participant factor on the product ratings on 17 different attributes using Wilks’ $\Lambda$. Subsequently, we examined differences between carrots for each attribute separately. In accordance with Stevens (2002), in case the Mauchly test was significant indicating a violation of sphericity and the Greenhouse-Geisser coefficient $\varepsilon \leq 0.7$, the respective $F$-Test was performed with the Greenhouse-Geisser correction. We averaged the $\varepsilon$ values from Greenhouse-Geisser and Huynh-Feldt when $\varepsilon > 0.7$. If the $F$-test revealed a significant difference between carrots, post hoc analysis was conducted with Bonferroni correction for pairwise comparisons. Furthermore, we performed a repeated measures ANOVA using preparation mode (5 levels: cooked, mashed, soup, baked, or raw) and carrot type (9 levels) as within-participant factors. Again, Bonferroni correction was used in paired comparisons. To investigate specific interrelationships between responses on the different response scales, we calculated Pearson correlation coefficients, using the individual judgments of each stimulus as replicates ($N = 40$ participants $\times 9$ stimuli $= 360$). All statistical analyses were performed with IBM SPSS version 22.

As regards the spontaneous associations, we counted the number of times associations to fruits and vegetables (e.g., carrot, tomato, turnip) and adjectives (e.g., artificial, fresh, sweet) were mentioned. In addition, references to production (e.g., grandmother’s garden, farm, genetically modified) were counted as they may suggest links with naturalness or artificialness. Other types of associations included kitchen tools (e.g., blender, potato cutter), points of purchase (e.g., supermarket) and food applications (e.g., soup, snack), but these were regarded as less relevant and are not reported. The number of associations is reported in parentheses in the next sections.

4. Results

In doubly multivariate analysis of variance in which we included all nine carrots and the responses on the 17 attributes, we found a highly significant effect of the type of carrot on responses [Wilks’ $\Lambda = 0.002$; $F(17,23) = 651.1, p < 0.001, \eta^2 = 0.998$]. This multivariate effect is supported by significant carrot effects for 16 attributes [all $F$'s $> 3.1, p$'s $< 0.01$]; only the effect for flexibility was not found to be statistically significant on the $p = 0.05$ level [$F(5.6,7.9) = 2.2, p = 0.051$].

We will now discuss analyses on the attributes separately for the four orange carrots (low, medium and high saturation, and brown spots) and for carrots with different hues (white-green, white, yellow, purple, red, and orange). The orange carrots at medium saturation level are used as the standard to which all others are compared, using paired comparisons with Bonferroni correction including nine carrots.

4.1. Orange carrots

To investigate the effect of color saturation, we first focus on responses for the three pictures of orange carrots with different saturation levels. Paired comparisons showed that carrots with low color saturation were rated significantly less attractive and fresh than the medium saturation carrots, whereas carrots with high color saturation were rated as significantly more artificial than the medium saturation level [all $p < 0.05$]. Possibly, an intermediate level of saturation is desirable, because it consistently obtains the highest mean ratings on attributes with a positive valence, such as attractiveness and purchase intention.
To investigate the effect of browning, we compared the responses for orange carrots with brown spots to the orange carrots with medium saturation. With browning, carrots are rated less attractive, healthy, and fresh than the evenly colored orange carrots. In addition, the purchase intention is lower for the carrots with brown spots ([all $p < 0.05$]) (Fig. 4).

In the spontaneous accounts, the medium saturation carrots were most often described as normal (7), crunchy (2), fresh (2) and healthy (2). The carrots with low color saturation were also described as normal (10) and healthy (2), but were also often considered natural (6) and old (5). The carrots with high saturation were less often called normal (4) or natural (1), but more often described as artificial (10), with some participants referring to dye, crayons, and Photoshop. The carrots with brown spots were also considered normal (6) and natural (5), but were clearly identified as old (7), rotting (4), or affected with disease or infestation (11).

### 4.2. Atypical colors

All five carrots with atypical colors are rated significantly lower on healthiness, attractiveness, purchase intention, familiarity, and similarity to known carrots ([all $p < 0.05$]) than the orange carrots. For the other attributes, the red carrots appear to follow the ratings of the orange carrots most closely, while the yellow and purple carrots also obtain quite similar response profiles. The white and white-green carrots show the largest number of differences with the orange carrots.

More specifically, the red carrots only deviate from the orange ones with higher ratings on sourness and spiciness ([all $p < 0.05$]). In the spontaneous associations, red carrots were sometimes described as normal (2), but also as genetically modified (2). Red carrots were mainly associated with radish (7), beetroot (7), chili (4), turnip (2), and tomatoes (2). These associations with other vegetables may explain why the red carrots were often expected to be rather spicy.

The purple carrots deviated from the orange carrots as they were rated less fresh and nutritious, and more bitter and artificial ([all $p < 0.05$]). Spontaneous reactions to the purple carrots were not so explicit, although some referred to their artificial or fake character (2), genetic modification (1) and the use of chemicals (1). Vegetable associations mostly involved beetroots (14), red onions (4), radish (3), and eggplant (1).

Yellow carrots were rated less nutritious, more sour and more artificial than the orange carrots ([all $p < 0.05$]). Spontaneous descriptors referred to their artificial character (5), with also references to genetic modification (3), nuclear waste (1), food coloring (2) and crayons (1). Yellow carrots were mostly associated with yellow bell peppers (5), lemons (2), pumpkin (1), butternut squash (1), and melon (1).

The responses for white and white-green carrots deviated most from the orange carrots. They both were rated lower in sweetness and higher in sourness, bitterness, and spiciness. They were considered to be less ripe and less nutritious. In addition, the white carrot is rated less fresh and more artificial, and the white-green carrot is rated less juicy than the orange carrot ([all $p < 0.05$]). Spontaneous descriptions for the white carrots included old (3) and rotten (1), unripe (2) and sour (1). Several respondents reported associations with winter, including cold and frozen (4 in total). Associations with vegetables included mainly intensely flavored products like radish (7), ginger (2), turnip (1), onion (1), fennel (1), and parsley (1). The white-green carrots were spontaneously described mainly as unripe (7) and sour (1). Similar to the white carrots, associations with other vegetables included many with strong flavors, such as radish (11), turnip (3), rutabaga (1), celery (3), leek (2), onion (1), cabbage (1), fennel (1), and parsley (1).

### 4.3. Preparation modes

In the repeated measures ANOVA with five preparation modes and nine carrot types, the within-subjects tests revealed significant main effects of Preparation mode [$F(4, 156) = 12.4, p < 0.001, \eta^2_p = 0.24$] and Carrot type [$F(5.5, 213.4) = 15.2, p < 0.001, \eta^2_p = 0.28$], and a significant Preparation mode × Carrot type interaction [$F(13.9, 541.1) = 4.3, p < 0.001, \eta^2_p = 0.10$]. This indicates that the way in which carrots are prepared and eaten partly depends on the type of carrot.

If we look at the grand means for the preparation modes, we can conclude that the likelihood was highest that participants would prepare and eat the carrots in cooked form (4.38), raw (3.94), or in soup (3.84), followed by mashed (3.32), and was lowest as baked (2.88). Mean ratings for cooked, raw, and soup are significantly higher than for baked, while ratings for cooked are also significantly higher than for mashed ([all $p < 0.05$]). The ordinal relationships between the means correspond to the mean amount of carrot (in g) the participants reported to consume per month for the different modes.
As regards differences between the types of carrots, participants were most likely to prepare and eat the orange carrots (4.68, 4.65, and 4.41 for low, medium and high saturation, respectively, and 3.65 for carrots with brown spots) or red carrots (3.80). They were less likely to prepare and eat the yellow (3.34), purple (3.19), white (2.85) or white-green (2.49) carrots. This pattern of mean ratings is comparable to the pattern found for purchase intention and attractiveness in Fig. 5. The means for the three orange carrots were significantly higher than the means for the yellow, purple, white and white-green carrots. The mean for the red carrots was significantly lower than for the orange carrots with low or medium color saturation, and was significantly higher than for the white and white-green carrots. The mean for the carrots with brown spots had a relatively large standard error and did not differ significantly from any of the other carrots ($p < 0.05$).

As is evident from the significant interaction, the consumption patterns differ somewhat between carrot types. Fig. 6 shows that the consumption likelihood generally decreases approximately monotonically from cooked, to soup, mashed and baked for all carrot types. However, eating raw carrots is relatively popular for some types (orange carrots with medium or high saturation), whereas it is relatively unlikely for several other types (orange carrots with low saturation or brown spots, and purple, white and white-green carrots).

The likelihood that carrots are prepared in a certain way could be related to the characteristics of the carrots. For each variable indicating the probability that participants would prepare and eat the carrot in a specific mode (raw, cooked, mashed, in a soup, baked) we calculated the Pearson correlation coefficients with the ratings that described how the carrots were perceived and evaluated. These analyses indicated that carrots were more likely to be eaten raw if they rated high on attractive ($r = 0.611$), healthy ($r = 0.486$), sweet ($r = 0.429$), nutritious ($r = 0.413$), and crunchy ($r = 0.332$), whereas carrots to be eaten cooked also needed to be rated healthy ($r = 0.494$), nutritious ($r = 0.452$), and sweet ($r = 0.358$), but overall attractiveness ($r = 0.395$) and crunchiness ($r = 0.220$) seemed less important. For carrots in soup, prepared mashed or baked, the correlation patterns were less pronounced, with highest correlations typically for attractive, healthy, and nutritious ($0.250 < \text{all } r's < 0.352$). All the mentioned correlations were statistically significant ($p < 0.001$).

5. Discussion

In a study conducted with photographic images of real carrots as they are available in shops, we investigated the expectations that differently colored products elicit among consumers. Because our study was motivated by a desire to optimize ecological validity and to study products in their everyday context, the factor of color may have been partly confounded with other factors, such as carrot size, diameter, shape, and so on. In addition, our color analysis showed that the hue of the orange carrots in the pictures differed marginally across variants and, thereby, may have slightly interfered with our assessment of the effect of differences in orange color saturation. These factors reflect a certain degree of natural product variation that is available in diverse buying situations, including supermarkets, outdoor markets and wholesale stores. This means that the present study gives valuable insights of the impact of color, but the exact amount of explained variance will be, as in all ecologically-driven studies, not precisely estimable. We decided for this approach as we mainly aimed for ecological and not for internal validity. We think that our results provide an interesting insight in the processes by which people form expectations of natural products with devaluing colors in everyday settings. Nonetheless, specific findings will need to be corroborated by replication studies that evaluate similar products in equivalent or different settings.

If we come back to the study results, we can see that we expected higher saturation to be associated with higher liking ratings for orange carrots, but we only found this for the medium saturation level compared to the low saturation level. For the highest saturation level, participants evaluated these carrots as more artificial than the carrots at medium level. We also predicted that people would expect carrots with more saturated orange colors to have a higher taste intensity, but the means for taste intensity were almost identical for the three types of orange carrots (Fig. 4). Possibly, the relationship between taste intensity and color is not related to the saturation dimension of color vision, but to the lightness dimension, with darker colors relating to higher taste intensity. It is also possible that this relationship mainly exists for processed products in which ingredients (e.g., fruit parts) are added to a neutral base (e.g., water or milk), rather than for agricultural products.

As concerns the effect of browning, our results clearly confirm our hypothesis that browning decreases attractiveness and purchase intention ratings. However, the likelihood that people will prepare and eat the carrots with brown spots is still relatively high. This suggests that people are unlikely to buy them, but will nevertheless cook them if they find them in their home storage. Interestingly, further inspection of responses in the individual preparation modes shows that orange carrots with brown spots and with low saturation are unlikely to be consumed raw. As these carrots tend to be described as old, this suggests that raw consumption is mainly reserved for fresh vegetables, while older vegetables may still be consumed after having undergone heat treatments. Our observations are in line with de Hooge et al. (2017), who found that an apple with a spot is more likely to be used in cooking than an intact apple.

Péneau, Brockhoff, Escher, and Nueslli (2007) used an expert panel to generate descriptors characterizing the freshness of carrots. Their panel suggested that the freshness of a carrot is characterized visually by a shiny surface that has no dark bruises, no film and is not shriveled. In addition, the carrot should not have any dried ends. In subsequent analyses of experimental data generated by a descriptive panel, they found that freshness ratings were positively related with shininess (+1.3) and film (+0.2), and negatively related with shrivel (−1.3) and color intensity (−0.8). It is not clear from their analysis whether the color assessments refer to the occurrence of incidental brown spots or could also involve changes in orange darkness or saturation. However, their study indicates that consumers use multiple visual variables, including color, as indicators of product freshness.

By analyzing and manipulating square images of small areas of surfaces of fruits and vegetables, Wada et al. (2010) showed that they could use information from simple luminance histogram statistics for an accurate estimation of freshness perception, rather than analyzing the color and deformation information typically used in computer algorithms that evaluate product quality (e.g., Sun, 2008). This approach has proven successful in analyzing images of carrot surfaces as well.
Furthermore, presenting participants with a large diversity of fresh food products in the subsequent study with product images would be very challenging, given that regional availability will be limited due to translucency as well (Fleming & Bulthoff, 2005). In the case of carrots, however, Arce-Lopera et al. (2015) only found significant correlations between product colors and ratings for sweetness (\(r = 0.348\)), bitterness (\(r = -0.267\)), sourness (\(r = -0.245\)), and nutritional value (\(r = 0.216\)) [all \(p < 0.001\)].

As concerns preparation modes, we found particularly strong correlations for carrots to be eaten raw on many positive attributes, while such correlations were weaker for carrots to be eaten cooked, and even weaker for the alternative preparation modes (mashed, baked, soup). This suggests that people are particularly picky regarding the carrots that they will consume in a raw state: these should be the most attractive, healthy, and crunchy. For the carrots that are cooked, people may be less demanding, also because an attribute like crunchiness may be less relevant for cooked carrots compared to raw tasting. For carrots in soup, prepared mashed or baked the correlation patterns were even less pronounced. However, for these products the low correlations may also be partly related to low consumption quantities, as these preparation modes were not so popular in the employed sample.

5.1. Limitations

In the present study, we used images of real products presented on a calibrated true color screen. This focus was purposefully used to test the impact of visual characteristics on the assessment of real-life products. Nonetheless, evaluations might differ when consumers are confronted with real products in shops where they can acquire further sensory information, such as olfactory or haptic cues. As freshness and ripeness can change very fast with natural products, product evaluations are likely to vary accordingly, due to changes in multiple sensory qualities. Hence, presenting real products will make the effects of color differences even harder to separate from any confounding factors. Furthermore, presenting participants with a large diversity of fresh food products that are similar in shape and differ only in color is likely to be very challenging, given that regional availability will be limited due to geographical and seasonal factors.

Although we strived for similar product shapes in the pictures, in a subsequent study with product images the product presentation could be even more systematic, for instance by photographing all stimuli with the same background and lighting, at the same degree of inclination, using more similar carrot shapes, and identical picture sizes. Although these variables were not the focus of our study, variation on these variables may have affected our outcomes. An alternative strategy to reduce the possible effects of uncontrolled differences could be to present participants with multiple pictures of carrots of a particular color or saturation level. For example, participants could evaluate three different pictures of red carrots, three different pictures of white carrots, and so on. If all three pictures were evaluated in similar ways and would differ between carrot types, that would add support for concluding that such differences were systematic and due to differences between product colors.

Other studies have used images of only small parts of surfaces of carrots. These studies have shown that analyzing such images can be used, for instance, to evaluate the freshness of carrots (Arce-Lopera et al., 2015). Using images of small pieces of surface has the advantage that they only contain information on the surface of the carrot and are thus more standardized. Participants who evaluate such images are not distracted by variations in background color, shape and orientation of the carrot, or the presence or absence of green leaves at the top of the carrot. On the other hand, by isolating the carrot surface from its context, the task is less natural, and the judgments may be of limited value in understanding consumers’ evaluations of carrots in real life.

Our participants were international students, which made the population quite heterogeneous in terms of cultural background. This heterogeneity may have added to the richness of our results, but may also have increased the variability of our dependent measures and thus reduced the amount of clear differences between product types. It would be interesting to see to what extent the results would differ if we could compare groups with different cultural backgrounds, because associations may differ depending on previous experiences. This could also partly solve the issue that in our study many of the participants actually had no experience with tasting the carrots displayed, whereas direct sensory experience is extremely important for determining the acceptance of novel foods (e.g., Koza, Cilmi, Dolesse, & Zellner, 2005; Tan, Fischer, van Trijp, & Stieger, 2016). Participants who have tasted the depicted carrots will probably generate very different responses for carrots that evoke incorrect associations (e.g., red carrot – red pepper, white carrot – horseradish).

5.2. Practical implications

To evaluate the effect of colors on the perception of foods, we studied carrots in two experiments. Our results are likely to apply to other types of fruits or vegetables as well. For instance, color changes that result from ripening (e.g., increased saturation and darkness) and deterioration (decreased saturation, more spots) will apply to all fruits and vegetables, just as unfamiliar colors are likely to have an important impact on expected taste and quality aspects. In both cases, insights in the cognitive processes that underlie consumers’ expectations can help to predict the effects of color differences. In the first case, knowledge of the characteristics of gradual ripening and decay processes for the focus product is key to understand consumer expectations of quality. In the second case, the present study has shown that spontaneous associations to other fruits and vegetables can generate an impression of the expected product properties.

To what extent will the current findings also apply to other food categories, such as dairy products or highly processed foods? Some products in the dairy category also show natural ripening processes, such as fermentations in the case of yogurt and cheese. In addition, all food products are susceptible to natural decay process, which may be microbiological, enzymatic, chemical, or physical in nature. Hence, for all food products knowledge of these naturally occurring changes can predict what consumers will expect when they perceive the associated color changes in the product. Similarly, for many food products...
unfamiliar colors will generate associations to other products, which are likely to lead to confusion and biased expectations that will affect the evaluation of the product during consumption (e.g., Schifferstein, 2001). Hence, in this case knowing the associations that product evokes is also likely to yield insight into what expectations consumers will generate.

Garber, Hyatt, and Starr (2001) have suggested that introducing foods with novel colors on the market may be a fruitful strategy for its potential to draw attention and differentiate between brands. Indeed, the large color variety of carrots provided nowadays opens up possibilities to increase the liveness of home and restaurant meals. However, consumers may be reluctant to buy these unfamiliar varieties in stores, as their expectations may be unsure and biased due to associations with unrelated vegetables. Especially in the case of white and white-green carrots, expectations were mostly negative and some familiarization activity seems necessary to increase acceptance.

For the realization of product acceptance among adult consumers, a step by step intervention may be most successful. A completely novel and unusual food requires accurate and accessible information to increase the probability to be tried and accepted (Cardello, Maller, Masor, Dubose, & Edelman, 1985). Hence, breeders, traders, and retailers may need to invest in providing information or creating tasting opportunities for consumers, in order to familiarize them with the sensory properties and preparation opportunities for these vegetables. In addition, grocery stores may stimulate consumption by offering combination packages containing mostly familiar orange carrots and some novel, purple or white carrots. By offering a possibility to try out the new varieties, consumers’ primary aversion can be gradually overcome. Possibly, in some cases the type of lighting used in retail outlets might be instrumental in improving acceptance of vegetables with atypical colors. For instance, Yang, Cho, and Seo (2016) found that light might be instrumental in improving acceptance of vegetables with new varieties, consumers novel, purple or white carrots. By omission, grocery stores may stimulate consumption by offering properties and preparation opportunities for these vegetables. In addition, grocery stores may stimulate consumption by offering combination packages containing mostly familiar orange carrots and some novel, purple or white carrots. By offering a possibility to try out the new varieties, consumers’ primary aversion can be gradually overcome.

Possibly, in some cases the type of lighting used in retail outlets might be instrumental in improving acceptance of vegetables with atypical colors. For instance, Yang, Cho, and Seo (2016) found that light can modulate consumers’ willingness to eat and their hedonic impressions of foods, such as apples and bell peppers. The light variations that are beneficial seem limited, as food samples presented under white or yellow light are appreciated more than samples presented under blue, red and green light. Nonetheless, subtler light changes in the white-yellow spectrum might be able to increase appeal, as they could mask subtle color differences between products with bright versus dull colors, such as the levels of orange saturation we tested. However, they are unlikely to mask the large color differences that indicate very different product specificities.

Some studies in the field of aesthetics have shown that people have a preference for the typical and the familiar (e.g., Verzyer & Hutchinson, 1998), whereas other studies have shown that people are attracted by the new, untypical and innovative (e.g., Carbon & Leder, 2005; Simonson & Nowlis, 2000). Although preference for typical products seems to be incompatible with a desire for the new, Hekkert, Snelders, and van Wieringen (2003) showed that the two tendencies come together in the MAYA design principle, striving for most advanced, yet acceptable new products. Hence, although carrots as such are familiar to most people, new tastes, new colors, or new shapes add the advancement aspects. Carefully balancing new and familiar aspects is key to successful food introductions (Pliner & Stallberg-White, 2000; Stallberg-White & Pliner, 1999). By systematic advertisement, novelty effects can turn into trendsetting products through mere exposure (Faerber, Leder, Gerger, & Carbon, 2010; Zajonc, 1968) and even more clearly through adaptation (Faerber & Carbon, 2013). Moreover, Carbon (2010) has shown that people adapt very fast - within a single experimental session – to a product with a new form language. Even after such a short period of time, preferences that were established and settled for a long time can shift towards the new designs (Carbon & Hesslinger, 2014). Therefore, despite the low familiarity and attractiveness ratings in Fig. 5, these studies suggest that people may begin to appreciate carrots with unfamiliar colors after very limited exposure.

Vegetables with novel colors or combinations of differently colored vegetables could also provide new opportunities for families with children. Introducing vegetables in new colors might help to overcome children’s dislike of particular vegetables. For instance, for children who do not like broccoli, introducing the vegetable in the color of one of their favorite vegetables might help them to overcome a threshold of trying the vegetable. Alternatively, hiding a disliked vegetable in a colorful assortment of unknown vegetables might activate a tendency to explore and try everything on the plate.

Interventions to increase the acceptance of novel products need to be evaluated in order for dieticians and food marketers to get a more accurate view on the success rate of their implementation. The present study can offer the first insights for increasing awareness about the variety of food products and their consumption expectations, and may be used to create practical interventions which lead to higher acceptance of novel and unusual food products.

Acknowledgement

The authors would like to thank the photographers and copyright owners of the carrot pictures for enabling us to reproduce their pictures in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodqual.2018.10.002.

References


