Consumers’ Purchase Decisions for Products with Nutrition and Health Claims
What Role Do Product Category and Gaze Duration on Claims Play?
Steinhauser, Johann; Janssen, Meike; Hamm, Ulrich

Document Version
Accepted author manuscript

Published in:
Appetite

DOI:
10.1016/j.appet.2019.104337

Publication date:
2019

License
CC BY-NC-ND

Citation for published version (APA):

Link to publication in CBS Research Portal
Consumers’ Purchase Decisions for Products with Nutrition and Health Claims: What Role Do Product Category and Gaze Duration on Claims Play?

Johann Steinhauser, Meike Janssen, and Ulrich Hamm

Journal article (Accepted manuscript*)

Please cite this article as:

DOI: https://doi.org/10.1016/j.appet.2019.104337

* This version of the article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the publisher’s final version AKA Version of Record.

Uploaded to CBS Research Portal: August 2020

© 2019. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/
Consumers’ purchase decisions for products with nutrition and health claims: What role do product category and gaze duration on claims play?

Johann Steinhauser, Meike Janssen, Ulrich Hamm

PII: S0195-6663(19)30138-2
DOI: https://doi.org/10.1016/j.appet.2019.104337
Article Number: 104337
Reference: APPET 104337
To appear in: Appetite

Received Date: 31 January 2019
Revised Date: 20 June 2019
Accepted Date: 27 June 2019

Please cite this article as: Steinhauser J., Janssen M. & Hamm U., Consumers’ purchase decisions for products with nutrition and health claims: What role do product category and gaze duration on claims play?, Appetite (2019), doi: https://doi.org/10.1016/j.appet.2019.104337.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Title:

Consumers’ purchase decisions for products with nutrition and health claims: what role do product category and gaze duration on claims play?

Authors: Johann Steinhauser¹, Meike Janssen², Ulrich Hamm¹

¹ Department of Agricultural and Food Marketing, Faculty of Organic Agricultural Sciences, University of Kassel, Steinstr. 29, 37213 Witzenhausen, Germany
² Department of Management, Society and Communication, Copenhagen Business School, Dalgas Have 15, 2000 Frederiksberg, Denmark

Corresponding author: Johann Steinhauser. Tel: +49 5542 98-1239
E-mail addresses: j.steinhauser@uni-kassel.de; mj.msc@cbs.dk; hamm@uni-kassel.de

Keywords: health claims; nutrition claims; eye tracking; visual attention; consumer behavior; purchase decision
Abstract

Labeling food packages with nutrition and health claims is a widely used practice. This study aims to contribute to the literature by examining the gaze and purchase behavior of consumers regarding food products with nutrition and health claims. A close-to-realistic purchase situation with three-dimensional food packages with nutrition, health, and taste claims was simulated while the participants' eye movements were measured using head-mounted eye tracking glasses. In the purchase situation, two food categories with differing perceived healthiness were offered, orange juice and milk chocolate. In total, 156 consumers participated in this study which was undertaken in Germany. The findings indicate that each claim was noticed by at least 85% of the participants and health claims were looked at longer than nutrition or taste claims. Furthermore, when compared to other participants, the longer a participant looked at a specific claim, the more likely the participant was to purchase the respective product. Even though the product category had no effect on the gaze duration on claims, it affected the purchase behavior. Nutrition claims were preferred for orange juice while taste claims were preferred for milk chocolate. Health claims were preferred for neither. Marketers can benefit from this study, as it shows the gaze duration on claims influenced the purchase likelihood. Another important finding is that there are great differences between product categories regarding the type of claim consumers prefer.
Introduction

The use of nutrition and health claims as a tool to highlight health-related aspects of food products is a widely used practice in North America and Europe (Al-Ani, Devi, Eyles, Swinburn, & Vandevijvere, 2016; Hieke et al., 2016; Pravst & Kušar, 2015; Devi et al., 2014). These claims are short phrases printed on the front of food packages indicating the nutritional and health-related qualities of a food product. A nutrition claim states that a food is endowed with a certain beneficial nutritional characteristic. A health claim additionally states that this nutritional characteristic has a beneficial health effect on the body. The third category, as per EU Regulation No. 1924/2006, Art. 2, par. 2.4–2.6, is the risk reduction claim which states a reduction in the risk of developing a disease. Due to the infrequent use of risk reduction claims in European countries (Kaur et al., 2016; Pravst & Kušar, 2015) as well as internationally (Mayhew et al., 2016), this type of claim was not included in this present work.

Previous literature described nutrition and health claims as highly successful tools to promote sales (Nestle, 2007; Wansink, 2005). Most consumer research studies showed that nutrition and health claims had a positive effect on preferences and purchase behavior (Kaur, Scarborough, & Rayner, 2017). However, in recent years a number of studies have revealed that nutrition and health claims also led to negative evaluations and lower purchase intentions of products by consumers (Bialkova, Sasse, & Fenko, 2016; Aschemann-Witzel & Grunert, 2015; Orquin & Scholderer, 2015; van Buul & Brouns, 2015; Maubach, Hoek, & Mather, 2014; Lähteenmäki, 2013; Lähteenmäki et al., 2010) or to a decrease in purchases (Kiesel & Villas-Boas, 2013; Berning, Chouinard, & McCluskey, 2011). The discrepancy in the reported effects of nutrition and health claims, ranging from positive to negative, has been pointed out by researchers such as Kaur et al. (2017), Hieke et al. (2015), Bruschi, Teuber, and Dolgopolova (2015) or Lähteenmäki (2013).

In several studies, researchers suggested that the perceived healthiness of the product category determines the direction of the effect of nutrition and health claims on consumers’ evaluations and purchase behavior (Aschemann-Witzel & Grunert, 2017; Stanca, Grunert, & Lähteenmäki, 2017; Bialkova et al., 2016; Fenko, Kersten, & Bialkova, 2016; Masson, Debecquet, Fischler, & Merdji, 2016; Talati, Pettigrew, Dixon et al., 2016; Lähteenmäki, 2013). However, these studies are still in disagreement as to whether nutrition or health claims lead to higher purchases and better evaluations of
products which are perceived as healthy or unhealthy. In this context, some of these studies further analyzed the influence of taste claims and whether taste claims or nutrition and health claims had a greater influence on evaluations and purchase behavior. Compared to a nutrition or health claim, a taste claim simply refers to the taste of a food product, such as ‘great taste’ (Bialkova et al., 2016; Choi, Paek, & Whitehill King, 2012) and is not regulated by law.

One conclusion reached by researchers was that future research on nutrition and health claims should examine such claims on authentic packages in more realistic and natural settings (Kaur et al., 2017; Lähteenmäki, 2013; Hieke & Taylor, 2012). The effects of nutrition and health claims should be measured with actual behavior and not just self-reported preferences (van Buul & Brouns, 2015; Wills, Storcksdieck genannt Bonsmann, Kolka, & Grunert, 2012). As consumers may not look at the attributes of a food package, e.g. nutrition tables or health claims, in the way they report in questionnaires, researchers suggested investigating the actual visual attention of consumers on food products (Ares et al., 2013).

Based on a Stimulus-Organism-Response (SOR) paradigm, products labeled with different claims are stimuli with visual distinctiveness; hence, they may cause a bottom-up effect on the attention of the consumer, the so-called organism (Duerrschmid & Danner, 2018; van der Laan, Laura N., Hooge, Ridder, Viergever, & Smeets, 2015). Before consumers form a purchase decision, they usually look at the products; thus visual attention is the starting point of any subsequent behavior (Duerrschmid & Danner, 2018; Meyerding & Merz, 2018; Meißner, Musalem, & Huber, 2016). Eye tracking is an appropriate method to measure the gaze behavior of consumers (Talati, Pettigrew, Hughes et al., 2016; Abrams, Evans, & Duff, 2015). A considerable amount of research on visual attention on food labels has been conducted (Grebitus & Davis, 2017; Bialkova, Grunert, & van Trijp, 2013; Piqueras-Fiszman, Velasco, Salgado-Montejo, & Spence, 2013). Insight into the current state of eye tracking research regarding food packaging was given by a recent literature review (van Loo, Grebitus, Nayga, Verbeke, & Roosen, 2018). The authors stated, that within this area of eye tracking research, “only a few studies have focused specifically on food choice” (van Loo et al., 2018). Several other authors confirmed that more research is needed to find out how visual attention on labels such as nutrition and health claims influences a subsequent product choice (Peschel, Orquin, & Mueller Loose, 2019; Duerrschmid & Danner, 2018; Meyerding & Merz, 2018). This
gave reason to opt for a choice test – using real food packages labeled with claims – in combination with eye tracking.

The overall aim of the present work was to analyze the effects of claims on purchase behavior in a real-life shopping experiment by examining consumers’ gaze behavior regarding different product categories and claim types. The present study is innovative for two reasons. Firstly, it went beyond previous survey-based research on claims. Instead of relying on the assumption that study participants noticed a claim, the method of head-mounted eye tracking was capable of showing to what extent consumers looked at claims in a real-life shopping environment and in what way the visual attention on these claims affected the purchase decision. Secondly, the study investigated the influence of the perceived healthiness of product categories on the purchase decisions for products labeled with claims. This is a contribution to the existing research as it could help to explain the observed contradiction in the results of previous studies.

Therefore, the research questions were: (1) To what extent do consumers look at claims while shopping? (2) Do nutrition, health, and taste claims have an effect on the purchase decision? (3) Does gaze duration on claims have an effect on the purchase decision? (4) Regarding the analyses (1) to (3), what are the differences between product categories and claim types?

Literature Review

Previous research has shown that consumers did not acknowledge the nutritional composition of a specific food product, but tend to categorize a food product either as healthy or unhealthy (Larkin & Martin, 2016; Orquin & Scholderer, 2015; Belei, Geyskens, Goukens, Ramanathan, & Lemmink, 2012; Gravel et al., 2012). Due to this dichotomized perception of food by consumers, using the terms healthy and unhealthy for food products is common throughout the literature (Fenko et al., 2016; Talati, Pettigrew, Dixon et al., 2016; Bruschi et al., 2015; Choi et al., 2012; Lalor, Kennedy, & Wall, 2011).

Previous studies mostly tested nutrition and health claims on food perceived as healthy, whereas product categories perceived as unhealthy were not in the focus of research (Cornish, 2012). However, across different product categories and across different countries, the study findings showed that one third to half of the products which were considered of poor nutritional quality carry a nutrition or a health claim (Al-Ani et al.,
According to Art. 4 of EU Regulation No. 1924/2006, so-called nutrient profiles were to be established by 2009 to prevent the use of nutrition and health claims on food with poor nutritional quality. However, these nutrient profiles still have not been established (Kaur et al., 2016; Pravst & Kušar, 2015). Therefore, labeling nutrition and health claims on products which can be considered nutritionally poor is allowed in the European Union. Orquin and Scholderer (2015) remarked that nutrition and health claims on ‘unhealthy’ products should be further studied, especially to determine whether these claims can outweigh a food product’s nutritionally poorer composition, eventually misleading consumers.

Study results are very different regarding whether nutrition and health claims on ‘healthy’ (Bialkova et al., 2016; Fenko et al., 2016; Choi et al., 2012) or on ‘unhealthy’ food (Maubach et al., 2014; Gravel et al., 2012) lead to positive consumer evaluations or purchase intentions. It can be argued that nutrition and health claims on ‘unhealthy’ food – a so-called mismatch – lead to positive preferences because food perceived as unhealthy can potentially benefit from highlighting the health aspects of the food (Bech-Larsen & Grunert, 2003; Kähkönen, Tuorila, & Lawless, 1997). Labeling nutrition and health claims on ‘unhealthy’ food seemed to be more reasonable for consumers than labeling them on ‘healthy’ food because that food is already perceived as healthy (Krutulyte et al., 2011). Moreover, seeing ‘healthy’ food with these claims might trigger consumers to question why a product which is already ‘healthy’ needs to be labeled with a nutrition or health claim (Lähteenmäki, 2013). Several studies showed that consumers preferred a nutritional enhancement in ‘unhealthy’ food because it reduced consumer’s guilt for eating unhealthily (Cornish, 2012; Lampila, van Lieshout, Gremmen, & Lähteenmäki, 2009). Thus, nutrition and health claims on ‘unhealthy’ food can act as a form of justification (Belei et al., 2012). On the other hand, advertising a food product with its strength – a so-called match-up – might lead to positive preferences due to the synergetic effects between a claim and a food product, i.e., a ‘healthy’ food benefits from relating it to healthiness through a health claim (Choi et al., 2012). Also, studies showed that consumers saw ‘healthy’ foods as more acceptable and even more credible carriers for health claims than ‘unhealthy’ foods, while nutrition and health claims on ‘unhealthy’ foods might induce skepticism and distrust (Lalor et al., 2011; Siró, Kápolna, Kápolna, & Lugasi, 2008).

The interplay of match-up and mismatch effects is not limited to nutrition and health claims but extends to taste claims and ‘unhealthy’ food. The taste of food is generally
among the most important decisional aspects for consumers’ purchase decisions (Fenko et al., 2016; Bruschi et al., 2015). Research showed that taste claims on ‘unhealthy’ products (Fenko et al., 2016; Choi et al., 2012) and ‘healthy’ products (Kim, Cheong, & Zheng, 2009) both lead to higher preferences. On the contrary, nutrition and health claims lead to a loss in perceived tastiness (Liem, Toraman Aydin, & Zandstra, 2012; Sabbe, Verbeke, Deliza, Matta, & van Damme, 2009). Simply calling a product ‘healthy’ already had negative effects on its anticipated pleasantness (Wardle & Huon, 2000). If consumers longed for ‘unhealthy’ food, they explained that they only cared about taste, while issues such as health took a backseat (Chan, Patch, & Williams, 2005; Balasubramanian & Cole, 2002) and they did not want to see nutritional modifications in ‘unhealthy’ food (Patch, Tapsell, & Williams, 2005). In other words, research showed that consumers were unwilling to compromise taste for health (Verbeke, 2006). Thus, it can be assumed that, especially for ‘unhealthy’ food whose mere purpose of consumption is pleasure, the taste is so important that nutrition and health claims which signal less tastiness could have a negative effect on its preferability (Berning et al., 2011). Further research is needed to understand the influence of the perceived healthiness of food products on the effect of nutrition and health claims (Choi et al., 2012).

Methods

Overview of the study and its mixed-method approach

A combination of a purchase simulation with head-mounted eye tracking glasses and a questionnaire was used to achieve the research aims. The purchase simulation allowed for the analysis of the effect of different claims on consumers’ shopping and purchase behavior. In what way the different claim types have an influence herein was addressed by rotating the claims on the products. Additionally, the products were from two product categories with different perceived healthiness to check for the influence of the category. Up to this point, any measured effects would be based on the assumption that the claims were responsible because the participants had seen the claims, as is common practice in research. However, the use of eye tracking devices in a purchase simulation overcomes the limitation of assuming that participants looked at the stimuli (Meyerding & Merz, 2018).

Generally speaking, eye tracking is a method for collecting data about the movements of the eye. While registering the participant’s eye movements, the eye tracking system also records the participant’s visual field. The outcome is a video for each participant in
which their recorded visual field is overlaid with their eye movements. The result of this recording is the information about where the participant looked. This method of eye tracking works because humans only see 2% of their visual field sharply. So, to acquire information from an object such as a food package, consumers must purposefully move their eyes with great frequency (Balcombe, Fraser, & McSorley, 2015). Visual attention is a good indicator of what information is acquired and most likely processed (Ares, Mawad, Giménez, & Maiche, 2014). However, looking at a certain object does not necessarily mean elaborating on the object. Nonetheless, there is a close relationship between gaze and mind because consumers mostly process the information which they are looking at in that specific moment (Duchowski, 2007). Thus, eye tracking measures visual attention, which sheds light on how consumers cognitively process visual information.

By recording participants’ eye movements, it was possible to examine whether consumers actually noticed the nutrition and health claims on the package fronts and whether looking at them had an effect on the participant’s behavior. Head-mounted eye tracking glasses allowed to further research the effects of nutrition and health claims in a close-to-realistic shopping environment. With these glasses, the participants were able to move freely in front of shopping shelves and look at the packages from different angles, take them off the shelf or turn them over to read further information on their sides. Showing frontal photos of the package on a computer monitor is a limitation of previous research and was acknowledged as such (Piqueras-Fiszman et al., 2013). To the authors’ knowledge, to date no study has used head-mounted eye tracking glasses in a close-to-realistic shopping environment and measured the effects of nutrition and health claims on consumer behavior.

Eye tracking data can answer questions such as where consumers looked and for how long but not those questions about the underlying reasons behind the gaze behavior (Duerrschmid & Danner, 2018). Therefore, a questionnaire was part of the experiment following the purchase simulation with the eye tracking glasses. This is seen as a promising combination of methods to gain greater insight into consumer behavior (Holmqvist et al., 2011). The self-administered questionnaire, which participants filled out on a computer, comprised the following variables measured on 7-point Likert scales. Perceived healthiness of the product category, of the offered brands and the offered brands in comparison to familiar brands (1 = very unhealthy; 7 = very healthy) were adapted from Ares, Giménez, and Gámbaro (2009). On a scale from 1 = very unimportant to 7 =
very important, participants indicated to what extent certain product attributes are important during the everyday purchase of orange juice and milk chocolate; adapted from Bruschi et al. (2015). After defining nutrition and health claims in an easily understood way, the participants were asked whether they pay attention to nutrition and health claims (either on ‘healthy’ or on ‘unhealthy’ food) on a scale from 1 = strongly disagree to 7 = strongly agree. Similarly, the belief in the claimed health benefit of the offered health claims was measured; item adapted from Singer, Williams, Ridges, Murray, and McMahon (2006). The trust in the shown nutrition and health claims was measured on a scale from 1 = very untrustworthy to 7 = very trustworthy; adapted from van Herpen and van Trijp (2011). Finally, a few socio-demographic questions were asked.

Study design and data collection

Participants

All participants were recruited in the pedestrian area of a German city’s main shopping promenade. The medium-sized German city of Kassel (199,062 inhabitants) has an average population with average purchase power (age: 18–44: 49%, 44–80: 51%; household size: 1.9; monthly household net-income: €1821.5; Kassel - Department of Statistics (2018)). Recruiters stood at predefined spots where they were instructed to approach every third person passing by, resulting in a convenience sample. Participants were recruited on every day of the week and during the entire day to ensure a representative sample of shoppers. After the recruiter approached a person they asked if the person wanted to take part in a study about food. After completion of the task, participants received a remuneration of 10 euro. If the person declined, the recruiter asked why and wrote down the reason given as well as the assumed sex and estimated age of the person. If the person was interested in taking part, they were asked two screening questions; did they go grocery shopping at least occasionally and did they purchase orange juice and chocolate at least occasionally. If the person replied positively, they were asked to follow the recruiter to the nearby university building in which the experiment took place. A total of 5,112 of the people approached declined participation (reason in descending order: lack of time, disinterest in the subject, generally unwilling to participate in studies, sickness, concern about data privacy) or deemed to be unsuitable for taking part in the experiment (reason in descending order: language difficulties, not purchasing the requested products, medical condition related to the requested products or their ingredients, severe medical condition related to eyesight). The recruiters were university students instructed not to reveal the study’s specific purpose and to avoid
discussions on certain topics such as consumer behavior, health claims or healthiness of food products. Due to the complexity of the experiment, it was conducted by two scientific assistants. For the analyses, a sample with 156 participants was used whose characteristics are displayed in Table 1.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
<th>Sample 1</th>
<th>Population country 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (N = 153)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;45 years</td>
<td></td>
<td>53.6 %</td>
<td>44.1 %</td>
</tr>
<tr>
<td>&gt;=45 years</td>
<td></td>
<td>46.4 %</td>
<td>55.9 %</td>
</tr>
<tr>
<td>18-24</td>
<td></td>
<td>17.0 %</td>
<td>9.8 %</td>
</tr>
<tr>
<td>25-44</td>
<td></td>
<td>36.6 %</td>
<td>31.6 %</td>
</tr>
<tr>
<td>45-64</td>
<td></td>
<td>34.6 %</td>
<td>38.2 %</td>
</tr>
<tr>
<td>65-80</td>
<td></td>
<td>11.8 %</td>
<td>20.4 %</td>
</tr>
<tr>
<td>∅</td>
<td></td>
<td>41.2 %</td>
<td>44.3 %</td>
</tr>
<tr>
<td><strong>Sex (N = 156)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td>49.0 %</td>
<td>51.2 %</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td>51.0 %</td>
<td>48.8 %</td>
</tr>
<tr>
<td><strong>Educational level (N = 156)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No school graduation</td>
<td></td>
<td>1.9 %</td>
<td>4.1 %</td>
</tr>
<tr>
<td>9 years of schooling</td>
<td></td>
<td>14.0 %</td>
<td>33.2 %</td>
</tr>
<tr>
<td>10 years of schooling</td>
<td></td>
<td>22.9 %</td>
<td>28.4 %</td>
</tr>
<tr>
<td>University-entrance qualification</td>
<td></td>
<td>61.1%</td>
<td>34.3 %</td>
</tr>
<tr>
<td><strong>Household size (N = 156)</strong></td>
<td>Number of household members</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>48.4 %</td>
<td>41.1 %</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>24.8 %</td>
<td>34.0 %</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>16.6 %</td>
<td>12.3 %</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>6.4 %</td>
<td>9.3 %</td>
</tr>
<tr>
<td>&gt;=5</td>
<td></td>
<td>3.8 %</td>
<td>3.4 %</td>
</tr>
<tr>
<td>∅</td>
<td></td>
<td>1.9</td>
<td>2.0</td>
</tr>
<tr>
<td><strong>Children (N = 156)</strong></td>
<td>Number of children in the household</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>73.9 %</td>
<td>71.7 %</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>15.3 %</td>
<td>14.6 %</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>7.6 %</td>
<td>10.3 %</td>
</tr>
<tr>
<td>&gt;=3</td>
<td></td>
<td>3.2 %</td>
<td>3.3 %</td>
</tr>
<tr>
<td>Households with children</td>
<td></td>
<td>26.1 %</td>
<td>28.3 %</td>
</tr>
<tr>
<td><strong>Income (N = 156)</strong></td>
<td>The net-income of the household</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 900 €</td>
<td></td>
<td>35.0 %</td>
<td>10.02 %</td>
</tr>
<tr>
<td>900 – 1500 €</td>
<td></td>
<td>17.8 %</td>
<td>18.97 %</td>
</tr>
<tr>
<td>1500 – 2600 €</td>
<td></td>
<td>23.6 %</td>
<td>31.28 %</td>
</tr>
<tr>
<td>2600 – 4500 €</td>
<td></td>
<td>15.3 %</td>
<td>26.94 %</td>
</tr>
<tr>
<td>4500 – 6000 €</td>
<td></td>
<td>5.1 %</td>
<td>12.79 %</td>
</tr>
<tr>
<td>&gt; 6000 €</td>
<td></td>
<td>2.5 %</td>
<td></td>
</tr>
</tbody>
</table>

Note:
1 Source: Based on the information participants gave in a self-administered computer assisted interview at the end of the experiment. Information about race or ethnicity was not collected.
2 Source: Destatis (2017): German population 18 years until 80 years of age in 2017; own calculations based on Federal Statistical Office Germany.

**Table 1: Socio-demographic characteristics of the sample**
Instruments and procedures

Eye movements were recorded using a head-mounted eye tracking device (SMI Eye Tracking Glasses 2 Wireless, Table 2) which recorded both of the participants’ eyes. A head-mounted eye tracking device is susceptible to a loss in recording quality due to environmental influences such as variations in lighting and fluctuating distances between stimuli and participant which are notable in an in-store environment. To avoid the eye tracking system losing track of the eyes and the occurrence of the parallax error (Mansouryar, Steil, Sugano, & Bulling, 2016; Narcizo & Hansen, 2015), the study was performed in a laboratory which permitted the semblance of a shopping experience within a controlled environment.

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>SensoMotoric Instruments GmbH, Germany</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>SMI ETG 2w</td>
</tr>
<tr>
<td>Human interface design</td>
<td>Non-invasive video-based glasses-type eye tracker</td>
</tr>
<tr>
<td>Calibration</td>
<td>1-/3-point calibration</td>
</tr>
<tr>
<td>Sampling rate</td>
<td>60 Hz binocular</td>
</tr>
<tr>
<td>Gaze tracking accuracy</td>
<td>0.5° over all distances</td>
</tr>
<tr>
<td>Gaze tracking range</td>
<td>80° horizontal, 60° vertical</td>
</tr>
<tr>
<td>Scene camera</td>
<td>Resolution: 1280x960p @24 fps; 960x720p @30 fps; HDR (high dynamic range) mode with high sensitivity for low light</td>
</tr>
<tr>
<td>Scene camera field of view</td>
<td>Field of view: 60° horizontal, 46° vertical</td>
</tr>
<tr>
<td>Eyewear compatibility</td>
<td>Works with contact lenses and most vision correction spectacles; Snap-on corrective lenses from +/- 4 diopter available</td>
</tr>
</tbody>
</table>

Table 2: Technical data of the SMI Eye Tracking Glasses 2 (SensoMotoric Instruments GmbH, 2019)

Two pilot tests with 16 and 18 participants respectively were conducted to check the eye tracking part of the experiment. Claims on food packages are relatively small objects to examine with head-mounted eye tracking systems, thus keeping the quality of the recorded eye tracking data high was of upmost importance. Several improvements were made, some of which are listed below. The lighting in the room had to be kept bright and stable, while sunlight had to be blocked out. The calibration of the eye tracking glasses to the participants’ individual eye characteristics (3-point) was performed in front of a shopping shelf with dish detergent (points were glued on the products) instead of a blank poster with only the calibration points printed on it. This ensured that the calibration was performed approximately at the same distance as the participants would naturally stand in front of the product shelves. This resulted in an enhanced eye tracking quality in the subsequent shopping task. To analyze gaze behavior on the level of indi-
individual attributes labeled on food packages, rectangular and solid packages proved to be better than round or baggy packages. Abrupt changes in distances (parallax error) and looking through the eye tracking glasses close to their edges led to a loss in recording quality. To minimize this, each product category was placed in an individual shopping shelf with the products at its center and at the same eye level.

After the participants entered the laboratory, they were briefly introduced to the experiment which was presented as a simple shopping task including eye tracking. The study’s specific purpose was not revealed. The eye tracking glasses were handed to the participants with the instructions to wear them as they would normal glasses. Any pre-existing eyesight problems of the participants were corrected by mounting SMI’s optical lenses on the eye tracking glasses. Then, the eye tracking glasses were calibrated to the participants’ individual eye characteristics. As soon as these requirements were satisfied, the interviewer continued with reading the task instructions to the participant:

Translation into English: Imagine that you are shopping now in a normal grocery store. Behind the next wall, you will find the grocery store in which you are going to shop and pay with your own money. You need these groceries: orange juice and chocolate. You buy one product each, thus one container of orange juice and one bar of chocolate. Choose the products you would choose in your normal shopping situation. Take as much time as you usually need. The shopping basket is to your right and here we go.

The stimuli were three-dimensional food packages placed on shopping shelves. Each participant was asked to purchase one product in each product category (orange juice and chocolate), for a total of two purchased products. Three alternatives were offered in each product category. The participants placed their purchased products in the provided shopping basket. After the participants finished their shopping, the eye tracking glasses were removed and the participants were seated in front of a computer to fill out the questionnaire. At the end, the participants were debriefed and given their remuneration. Besides the pilot tests solely for the eye tracking part of the experiment, one final pilot test including the whole experiment (eye tracking and questionnaire) was conducted with 14 participants. The comprehensibility of the instructions and the correct interpretation of the items were improved where necessary.
Stimuli

Tested product categories

Orange juice and milk chocolate were the tested product categories as they differ in their perceived healthiness. The reason for consuming chocolate is purely hedonic, while the health aspect is irrelevant (Di Monaco, Ollila, & Tuorila, 2005). Chocolate is seen and used by researchers as an ‘unhealthy’ food product category in their studies (Belei et al., 2012; Chernev, 2011; Lalor et al., 2011). The opposite applies to orange juice which is seen and used as a ‘healthy’ food product category by researchers (Chernev, 2011; Siró et al., 2008; Bech-Larsen & Grunert, 2003). Besides their differences in perceived healthiness, milk chocolate and orange juice are very familiar to many consumers.

Tested nutrition, health, and taste claims

As pointed out in previous studies, the familiarity of the ingredients mentioned in nutrition and health claims might influence the effect nutrition and health claims have on consumers’ evaluations (Lähteenmäki et al., 2010; Ares et al., 2009; Bech-Larsen & Scholderer, 2007). Consequently, any possible interferences had to be eliminated by using familiar ingredients in the claims. Research showed that consumers were very familiar with vitamin C and calcium (Masson et al., 2016; Krystallis & Chrysochou, 2012; Bech-Larsen & Scholderer, 2007) including German consumers (Bornkessel, Bröring, Omta, & van Trijp, 2014). Health claims about vitamins referring to a benefit to the immune system and health claims about calcium referring to a benefit to the bones were also among the most commonly used health claims in the EU (Hung, Grunert, Hoefkens, Hieke, & Verbeke, 2017). Therefore, vitamin C and calcium were used as ingredients for the claims.

Furthermore, study results showed that the carrier-ingredient fit could have an influence on the effects of nutrition and health claims on preferences (Aschemann-Witzel & Grunert, 2017). Research found that fruit is mainly associated with vitamins whereas dairy products are associated with calcium (Masson et al., 2016). To avoid inadvertent influences, a carrier-ingredient fit between the tested ingredients and the tested product categories was established: the claims about calcium were labeled on milk chocolates and the claims about vitamin C were labeled on orange juices. To avoid further inadvertent influences, all tested claims were framed positively (Lähteenmäki, 2013; Kahneman & Tversky, 1979).
An overview of the tested claims is given in Table 3. The tested nutrition claims were labeled on food products which contained the mentioned nutrient in a sufficient amount and therefore complied with the rules of EU Regulation No. 1924/2006 Art. 5 par. 1.b. and EU Regulation No. 1169/2011 annex XIII part A. The tested health claims were authorized for use by the EFSA as stated in the online EU Register of nutrition and health claims made on foods (EFSA, 2019), according to EU Regulation No. 1924/2006, Art. 10 par 1. Taste claims are not subject to regulation.

If neither a nutrition nor a health claim was labeled on the package, a taste claim was present. This is common practice in the research area of nutrition and health claim research (Wong et al., 2014; Choi et al., 2012; Aschemann & Hamm, 2009) and counters the mere label effect. The mere label effect is a positivity bias towards a product that occurs solely because of the presence of any label or claim on a package front (Andrews, Netemeyer, & Burton, 2009).

The three claims, nutrition, health, and taste were rotated among the three product alternatives in each product category, creating six choice sets per product category. Thus, in every choice set, one package carried a taste claim, one package carried a nutrition claim, and one package carried a health claim. Each participant was given one choice set for each of the two product categories and the choice sets were equally distributed among the participants.

<table>
<thead>
<tr>
<th>Orange juice</th>
<th>Milk chocolate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nutrition claim</td>
<td>Rich in vitamin C</td>
</tr>
<tr>
<td>Health claim</td>
<td>Vitamin C contributes to the normal function of the immune system</td>
</tr>
<tr>
<td>Taste claim</td>
<td>Simply delicious</td>
</tr>
</tbody>
</table>

**Table 3: Claims used in the study on the front of the packages**

**Package and shelf design**

To make the purchase simulation as realistic as possible, the product alternatives were adapted from existing product packages. Since grocery shopping is characterized by habitual processes, familiar brand names and packages might influence consumers’ choice and the way they look at the packages (Graham, Orquin, & Visschers, 2012; Pieters, Rosbergen, & Wedel, 1999). Therefore, brands unfamiliar to the participants were
tested which has also been done and is recommended by previous researchers (Peschel et al., 2019; Singer et al., 2006). Brands from other German-speaking countries were used, specifically store brands from Austria and Switzerland which were not sold in Germany. The packages were of average design and were typical for the food category. The claims were well-incorporated into the front package design to avoid participants noticing them and thus becoming more engaged than they would have been in a normal purchase situation (Orquin & Scholderer, 2015). The study of the effects of nutrition and health claims on the participant’s behavior was performed without any forced exposure to the claims which was a common practice in previous studies in this research area (Aschemann-Witzel & Hamm, 2010). The claims were written with a font size of at least 14. The surface size of the claims on each brand stayed the same irrespective of the type of the claim (nutrition, health or taste claim). Thus, a health claim did not span larger across the surface area of the product’s front than a nutrition or a taste claim. The presentation order of the three brands in each product category on the shelf was not rotated. To avoid any bias due to upper or lower shelf-placement, both product categories were placed in individual shopping shelves and on the same eye level. To more clearly illustrate, photographs of the two shelves with the products in one model choice set are shown in Figure 1 and Figure 2.
Figure 1: The shelf with orange juices showing one model choice set

Figure 2: The shelf with milk chocolates showing one model choice set
Besides differences in the graphical layout of the packages, all other attributes were identical. The product name (“100 % orange juice”; “milk chocolate”), package size, ingredient list, and all other mandatory information were matched. Following the introduction of EU Regulation No. 1924/2006 Art. 7, every nutrient referred to in a nutrition and health claim must be additionally included in the nutrition facts table with the amount present in the product. The nutrition facts tables were made identical across each product category and represented the usual amounts for these products. Any optional / marketing information such as certification labels or logos (e.g. UTZ, Rainforest Alliance) were either removed or matched along all products within one product category. Since the participants were told to purchase the groceries, the shopping shelves included the price tags in the label strips. The three different price levels rotated among the three product alternatives per category. The prices (orange juice: €1.09; €1.29; €1.49; milk chocolate: €0.59; €0.79; €0.99) are within the typical price range for these product categories in Germany, which were validated by an inventory in different shops.

**Data analysis methods**

To start with, univariate methods were used to show descriptive statistics about the participants’ gaze behavior at the areas of interest. The areas of interest are defined as regions or elements on the stimuli which are of importance for the present research (Holmqvist et al., 2011). Analyzing mean gaze durations and proportions of participants gazing at an area of interest is a typical approach in eye tracking research (Duerrschmid & Danner, 2018; Ares et al., 2014; Holmqvist et al., 2011; Jacob & Karn, 2003). After this, paired sample t-tests were applied to check for differences between claim types and product categories on mean gaze duration. The differences in the purchase frequencies of the products were analyzed with non-parametric chi-square tests. For a combined analysis of the different claim types and gaze duration on the purchase decision, two multinomial logistic regression models were calculated; one model for each product category.

The analyses were performed with the gaze duration on the respective area of interest, such as the package or claim, if not otherwise specified. The so-called ‘dwell time’ was used for the analyses. Dwell time is the sum of all durations from the fixations and saccades in a certain area of interest. Once the eyes of a participant gaze on a certain area of interest, the time starts counting and stops when the participants’ eyes leave this area of interest. The summary of all the visits’ durations is the dwell time. The dwell time,
also referred to as gaze duration or glance duration, is a common measure in eye tracking research (Duerrschmid & Danner, 2018; Holmqvist et al., 2011; Jacob & Karn, 2003).

## Results

First, the average gaze durations of participants on the packages and on the different claim types, as well as the average time of the purchase decision for each product category, are illustrated. Then the differences in purchase frequencies between the claim types and between the product categories are shown. Results from the questionnaire are then used to explain the differences in the results obtained regarding claim type and product category. Finally, whether the gaze duration on claims affected the purchase decision is analyzed.

### Gaze duration on claims and products

The gaze duration on the claims varied considerably across the participants (Table 4). Each claim was looked at by at least 85% of the participants.

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
<th>SE</th>
<th>Median</th>
<th>Share of participants with zero views</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Orange juice (N = 156)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taste claim</td>
<td>0 ms</td>
<td>4381 ms</td>
<td>946 ms</td>
<td>758 ms</td>
<td>60</td>
<td>796 ms</td>
<td>8.9%</td>
</tr>
<tr>
<td>Nutrition claim</td>
<td>0 ms</td>
<td>5775 ms</td>
<td>1162 ms</td>
<td>906 ms</td>
<td>72</td>
<td>946 ms</td>
<td>6.4%</td>
</tr>
<tr>
<td>Health claim</td>
<td>0 ms</td>
<td>4945 ms</td>
<td>1373 ms</td>
<td>1187 ms</td>
<td>94</td>
<td>1028 ms</td>
<td>5.7%</td>
</tr>
<tr>
<td><strong>Milk chocolate (N = 156)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taste claim</td>
<td>0 ms</td>
<td>3900 ms</td>
<td>1019 ms</td>
<td>795 ms</td>
<td>63</td>
<td>879 ms</td>
<td>12.7%</td>
</tr>
<tr>
<td>Nutrition claim</td>
<td>0 ms</td>
<td>3966 ms</td>
<td>1145 ms</td>
<td>855 ms</td>
<td>68</td>
<td>962 ms</td>
<td>7.6%</td>
</tr>
<tr>
<td>Health claim</td>
<td>0 ms</td>
<td>8049 ms</td>
<td>1543 ms</td>
<td>1533 ms</td>
<td>122</td>
<td>1111 ms</td>
<td>14.6%</td>
</tr>
</tbody>
</table>

*Table 4: Gaze durations on claims*

In both product categories, participants spent more time looking at health claims than they did the two other claim types. For orange juice, participants spent, on average, 0.95 seconds looking at taste claims, 1.16 seconds looking at nutrition claims, and 1.37 seconds looking at health claims. The gaze duration differences among the claim types were significant (paired sample t-tests). For milk chocolate, participants spent 1.02 seconds looking at the taste claim, 1.15 seconds at the nutrition claim, and 1.54 seconds at the health claim. Again, all differences were significant.

Between the two product categories, there was no significant difference in participants’ gaze duration on claim types, i.e. the gaze duration on, e.g. the nutrition claims, was the
same for orange juice and milk chocolate (Table 5). In both categories, participants spent around a third of their time looking at the claims in relation to the time spent on the package fronts. Considering the gaze duration on all the package sides together (orange juice: front, back, left, right; milk chocolate: front, back), there was no significant difference between the product categories in the amount of time participants looked at the packages; on average 15.79 seconds for orange juice and 16.72 seconds for milk chocolate. Including gaze duration on price tags and on objects other than the packages, e.g. the shelf itself, the whole purchase decision took, on average, 19.78 seconds for milk chocolate and 18.56 seconds for orange juice with no significant difference between the two. Further results show that participants looked significantly longer at the nutrition tables on orange juice than the ones on milk chocolate, but significantly shorter at the brand names as well as the brand logo on orange juice compared to milk chocolate (Table 5).

<table>
<thead>
<tr>
<th>Gaze durations on claims [in ms]</th>
<th>Orange juice Means (SD)</th>
<th>Milk chocolate Means (SD)</th>
<th>t value</th>
<th>Effect size r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taste claims</td>
<td>946 (758)</td>
<td>1019 (795)</td>
<td>-.9539</td>
<td>n.s.</td>
</tr>
<tr>
<td>Nutrition claims</td>
<td>1162 (906)</td>
<td>1145 (855)</td>
<td>.2028</td>
<td>n.s.</td>
</tr>
<tr>
<td>Health claims</td>
<td>1373 (1187)</td>
<td>1543 (1533)</td>
<td>-1.4257</td>
<td>n.s.</td>
</tr>
<tr>
<td>All 3 claims combined</td>
<td>3481 (2254)</td>
<td>3707 (2515)</td>
<td>-1.1186</td>
<td>n.s.</td>
</tr>
<tr>
<td>Aggregated gaze duration on the package fronts of the three products per category [in ms]</td>
<td>9764 (5691)</td>
<td>12298 (6930)</td>
<td>-5.5146 ***</td>
<td>.4050</td>
</tr>
<tr>
<td>Aggregated gaze duration on the price tags of the three products per category [in ms]</td>
<td>2714 (1916)</td>
<td>2902 (2209)</td>
<td>-1.2725</td>
<td>n.s.</td>
</tr>
<tr>
<td>Aggregated gaze duration on the nutrition tables of the three products per category [in ms]</td>
<td>2289 (5675)</td>
<td>1316 (4122)</td>
<td>2.6730 **</td>
<td>.2099</td>
</tr>
<tr>
<td>Aggregated gaze duration on the brand name of the three products per category [in ms]</td>
<td>1536 (1106)</td>
<td>2191 (1172)</td>
<td>-6.7362 ***</td>
<td>.4759</td>
</tr>
<tr>
<td>Aggregated gaze duration on the additional package sides of the three products per category (right + left + back for orange juice; back for milk chocolate) [in ms]</td>
<td>6022 (10449)</td>
<td>4420 (10205)</td>
<td>2.0221 *</td>
<td>.1603</td>
</tr>
<tr>
<td>Aggregated gaze duration on the whole packages of the three products per category [in ms]</td>
<td>15786 (13444)</td>
<td>16718 (14305)</td>
<td>-.9835</td>
<td>n.s.</td>
</tr>
<tr>
<td>Aggregated gaze duration during the whole purchase decision per category [in ms]</td>
<td>18562 (14189)</td>
<td>19775 (15415)</td>
<td>-1.2644</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Note: N = 156; Significance p < .001 = ***; p < .01 = **; p < .05 = *; p < .1 = *(*)
Effect of different claim types on purchase decision

Besides the effect on gaze duration, the different claim types also affected the purchase decision (Figure 3). The so-called expectancy value was 33.33% for a claim not influencing the purchase decision because the three claim types were equally present in each product set. For each claim type, whether the share of purchases was significantly different from the expectancy value was tested.

![Figure 3: Share of purchases by claim type](image)

In the product category ‘orange juice’, participants bought products labeled with the nutrition claim significantly more often (χ² (1) = 4.1407, p = .0419). The taste claim and the health claim had no significant effect, i.e. the share of purchases was not significantly different from the expectancy value (Table 6). In the category ‘milk chocolates’, participants bought products labeled with the taste claim significantly more often (χ² (1) = 3.4904, p = .0617). Products labeled with a health claim were significantly less preferred (χ² (1) = 4.8750, p = .0272). The nutrition claim did not show any effect compared to the expectancy value.

<table>
<thead>
<tr>
<th>Claim Type</th>
<th>Orange juice (N = 156)</th>
<th>Milk chocolate (N = 156)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taste</td>
<td>30.6%</td>
<td>34.6%</td>
</tr>
<tr>
<td>Nutrition</td>
<td>40.8%</td>
<td>25.0%</td>
</tr>
<tr>
<td>Health</td>
<td>28.7%</td>
<td>40.4%</td>
</tr>
<tr>
<td>Expectancy</td>
<td>33.33%</td>
<td>33.33%</td>
</tr>
<tr>
<td></td>
<td>Share of purchases</td>
<td>chi-square</td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Orange juice</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taste claim</td>
<td>30.57%</td>
<td>0.4601</td>
</tr>
<tr>
<td>Nutrition claim</td>
<td>40.76%</td>
<td>4.1407 *</td>
</tr>
<tr>
<td>Health claim</td>
<td>28.66%</td>
<td>1.4100</td>
</tr>
<tr>
<td><strong>Milk chocolate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taste claim</td>
<td>40.38%</td>
<td>3.4904 (*)</td>
</tr>
<tr>
<td>Nutrition claim</td>
<td>34.62%</td>
<td>0.1154</td>
</tr>
<tr>
<td>Health claim</td>
<td>25.00%</td>
<td></td>
</tr>
<tr>
<td><strong>Orange juice vs. milk chocolate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taste claim</td>
<td>30.57% vs. 40.38%</td>
<td>6.816 **</td>
</tr>
<tr>
<td>Nutrition claim</td>
<td>40.76% vs. 34.62%</td>
<td>2.832 (*)</td>
</tr>
<tr>
<td>Health claim</td>
<td>28.66% vs. 25.00%</td>
<td>1.231</td>
</tr>
</tbody>
</table>

Note: N = 156; Significance p < .001 = ***; p < .01 = **; p < .05 = *; p < .1 = (*)

**Table 6: Comparison of share of purchases by claim types and product categories (chi-square tests)**

There were also significant differences between the product categories. A taste claim on milk chocolate led to a significantly larger share of purchases than a taste claim on orange juice. Conversely, a nutrition claim on orange juice led to a larger share of purchases than a taste claim on milk chocolate. Regarding health claims, there was no significant difference in the share of purchases between the product categories.

**Participants attitudes towards the two product categories**

As shown so far, there was a difference in gaze durations between the two product categories as well as a difference in purchases in combination with different claim types. In line with a priori assumptions, there was a significant difference in consumer perception of the two product categories as the following results demonstrate. Participants perceived orange juice, on average, healthier than milk chocolate (Table 7). Likewise, participants perceived the offered brands of orange juices healthier than the offered brands of milk chocolates. Participants further said they paid more attention to nutrition and health claims labeled on ‘healthy food’ compared to ‘unhealthy food’. Thus, participants not only perceived orange juice as healthier, but also cared more about nutrition and health claims on orange juice. However, the actual purchase and gaze behavior of the participants did not fully correspond with the stated consumer perceptions. Participants’
gaze duration on the nutrition and health claim showed no difference between the two
categories and products labeled with the health claim were the least preferred for pur-
chase in both categories. In accordance with the stated perceptions, participants bought
orange juice more often when it was labeled with a nutrition claim.

<table>
<thead>
<tr>
<th>Means (SD)</th>
<th>t value</th>
<th>Effect size r</th>
</tr>
</thead>
</table>
| **I. Perceived healthiness of the product category**
Orange juice vs. milk chocolate | 4.68 (1.316) vs. 3.24 (1.469) | 9.671*** | .6122 |
| **II. Perceived healthiness of the offered brands**
Orange juice vs. milk chocolate | 4.09 (1.242) vs. 2.85 (1.249) | 10.179*** | .6317 |
| **III. Perceived healthiness of the offered brands in comparison to familiar brands**
Orange juice vs. milk chocolate | 3.69 (1.062) vs. 3.316 (1.061) | 3.766*** | .2887 |
| **IV. Paying attention to nutrition and health claims on healthy vs. unhealthy food**
“Healthy food” vs. “unhealthy food” | 3.91 (2.110) vs. 3.34 (2.139) | 3.830*** | .2932 |

Note: N = 156; Significance p < .001 = ***; p < .01 = **; p < .05 = *; p < .1 = (v)
7-point Likert scales; wording I: “How healthy do you think orange juice (milk chocolate) is on average?”
with 1 = very unhealthy, 7 = very healthy; II: “How healthy do you think the offered orange juice (milk chocolate)
brands are?” with 1 = very unhealthy, 7 = very healthy; III: “How healthy do you think the offered orange juice (milk chocolate)
brands you are familiar with?” with 1 = very unhealthy, 7 = very healthy; IV: “On healthy (unhealthy) food I
pay a lot of attention to nutrition & health claims” with 1 = strongly disagree, 7 = strongly agree.

Table 7: Consumer perceptions of the two product categories (paired sample t-tests)

Additionally, participants were asked how important the following attributes were when
shopping for orange juice and milk chocolate: Taste, price, healthiness, nutritional val-
ue, and brand. The means are shown in Figure 4. Taste was reported as being more
important for the purchase of milk chocolate than for orange juice. This corresponds
with the purchase behavior as taste claims on milk chocolate led to a significantly larger
share of purchases than on orange juice (Table 6). Price was rated the second most
important attribute in both categories, followed by healthiness, nutritional value, and
brand, which is reflected in the longer gaze durations on price tags in comparison to
nutrition tables or brand names (Table 5). For the following attributes, there is a differ-
ence in the ranking between the categories: Healthiness and nutritional value are more
important than the brand for the purchase of orange juice, whereas for milk chocolate,
the brand is more important than healthiness and nutritional value (Figure 4). This is in
accordance with the gaze duration as participants looked longer at nutrition tables than
at brand names on orange juices, while the opposite was true for milk chocolates (Table
5). The differences in ratings and gaze durations are also in accordance with the ob-
erved share of purchases by claim type and product category.
Note: The attributes were rated on a 7-point Likert scale, ranging from 1 = very unimportant to 7 = very important.

- The letter ‘a’ indicates a significant difference with the next attribute following in the rating within orange juice: taste – price – healthiness – nutritional value – brand (p < .05);
- The letter ‘b’ indicates a significant difference with the next attribute following in the rating within milk chocolate: taste – price – brand – healthiness – nutritional value (p < .05);
- The asterisk * indicates a significant difference between the two product categories within the same attribute (p < .05).

**Figure 4: Importance of product attributes within one product category and between the two product categories (paired sample t-tests)**

Participants were further asked to rate their level of trust in the claims tested in the purchase situation. The nutrition and health claims on orange juice were trusted significantly more than the respective claims on milk chocolate (Table 8). However, the low levels of trust in both categories show that participants were rather unsure whether to trust or distrust the claims. The belief in the claimed health benefit was significantly higher for health claims on orange juice than on milk chocolate. The lower levels of trust and belief towards claims on milk chocolate are in accordance with the share of purchases in that nutrition claims on milk chocolate led to a lower share of purchases than for orange juice (Table 6). In both categories, participants trusted the health claims more than the nutrition claims. However, this did not correspond with the purchase behavior since health claims led to the lowest share of purchases in both categories. The reason for this could be the low level of belief in the claimed health benefit: for orange juice, the participants on average neither agreed nor disagreed to believe in the claimed health benefit and for milk chocolate, they even disagreed to believe in the claimed health benefit.
Table 8: Consumer perceptions of the claims tested in the experiment (paired sample t-tests)

<table>
<thead>
<tr>
<th>Claims Tested</th>
<th>Orange Juice Means (SD)</th>
<th>Milk Chocolate Means (SD)</th>
<th>t value</th>
<th>Effect size r</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Trust in the shown nutrition claims</td>
<td>3.91 (1.642)</td>
<td>2.91 (1.638)</td>
<td>7.530 ***</td>
<td>.5163</td>
</tr>
<tr>
<td>II. Trust in the shown health claims</td>
<td>4.45 (1.820)</td>
<td>3.97 (2.115)</td>
<td>3.638 ***</td>
<td>.2797</td>
</tr>
<tr>
<td>III. Belief in the claimed health benefit of the offered health claims</td>
<td>3.54 (1.78)</td>
<td>2.12 (1.473)</td>
<td>10.307 ***</td>
<td>.6365</td>
</tr>
</tbody>
</table>

Note: N = 156; Significance p < .001 = ***; p < .01 = **; p < .05 = *; p < .1 = .(.)

7-point Likert scales; wording: I: “How trustworthy do you think the nutrition claim ‘rich in vitamin C’ on orange juice (‘rich in calcium’ on milk chocolate) is?” with 1 = very untrustworthy, 7 = very trustworthy; II: “How trustworthy do you think the health claim ‘Vitamin C contributes to the normal function of the immune system’ on orange juice (‘Calcium is needed for the maintenance of normal bones’ on milk chocolate) is?” with 1 = very untrustworthy, 7 = very trustworthy; III: “Imagine you are eating the offered orange juices (milk chocolates). Do you expect positive effects on the function of your immune system (on maintaining your bones)?” with 1 = strongly disagree, 7 = strongly agree.

Effect of gaze duration on claims on purchase decision

The results presented above show that the gaze durations on claims do not allow a conclusion towards the share of purchases to be reached; for instance, although health claims were looked at the longest (Table 4), orange juices and milk chocolates labeled with health claims were the least preferred in purchases (Table 6). For the analysis of a direct relationship between the participants’ individual gaze durations on claims and their purchase decision, multinomial logistic regression models were used. Separate models for each product category were calculated. The dependent variable was the purchase decision. Thus, the dependent variable had three categories: a product labeled with a nutrition claim, a health claim or a taste claim. The three independent variables were the gaze durations on each claim type.

For both models, multicollinearity of the three gaze variables was checked. The tolerance values were between 0.69 and 0.79 and the VIF values were between 1.32 and 1.45, thus far below the thresholds that indicate a multicollinearity problem (Urban & Mayerl, 2011; Menard, 2002). Furthermore, the variance proportion showed that no gaze variable had high proportions on the same eigenvalue, as this would indicate that the regression coefficients’ variances are dependent (Field, 2013).

Table 9 shows the parameter estimates of the two models. In both models, the reference category was ‘taste claim’. It was found that the longer a participant looked at a specific claim (nutrition, health or taste) compared to other participants, the more likely
the participant was to purchase the product with the respective claim. This relationship was significant across the three claim types and the two product categories.

<table>
<thead>
<tr>
<th>Coefficients B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model 1:</strong> Purchase of orange juice</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Gaze duration on nutrition claim</td>
</tr>
<tr>
<td>Gaze duration on health claim</td>
</tr>
<tr>
<td>Gaze duration on taste claim</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Gaze duration on nutrition claim</td>
</tr>
<tr>
<td>Gaze duration on health claim</td>
</tr>
<tr>
<td>Gaze duration on taste claim</td>
</tr>
</tbody>
</table>

Note: Significance p < .001 = ***; p < .01 = **; p < .05 = *; p < .1 = (*)

| **Model 1:** | | **Model 2:** |
| N = 156; R² = .1816 (Cox & Snell), .2049 (Nagelkerke). | | N = 156; R² = .1546 (Cox & Snell), .1748 (Nagelkerke). |
| Model χ²(6) = 31.4594, p = .00002 | | Model χ²(6) = 26.1988, p = .0002 |

**Table 9: Multinomial logistic regression models on claim types**

**Discussion and conclusions**

The purpose of this study was to investigate the role of product category and gaze duration in consumers’ purchase decisions for products with nutrition, health, and taste claims. The originality of this study lies in the fact that a combination of a close-to-realistic purchase simulation with eye tracking and a subsequent survey was used.

**Visual attention on claims**

Each claim was noticed by around 90% of the participants. However, the amount of time the participants looked at a claim differed significantly among the claim types with health claims being looked at the longest, followed by nutrition, and finally, by taste claims. The different lengths of the health, nutrition, and taste claims, with nine to ten, three, and two words respectively, as well as the different complexity in processing their information might have played a role in this result. When the participants’ gaze duration on an individual level was considered rather than sample means, it was found that the longer a participant looked at a certain claim, the more likely they were to purchase the respec-
tive product. This relation between a longer gaze duration on claims and a higher purchase likelihood was found for every claim type for both of the product categories. This result adds value to the existing research on nutrition and health claims as it shows, for the first time, a direct relationship between consumers looking at a claim on three-dimensional products and purchasing the respective product. The implication of this result is that a claim must capture the consumer’s attention so that they look at the claim longer, which eventually increases the likelihood of the product being purchased. To increase visual attention on a package label, the visual density of information, the so-called visual clutter, around the label should be decreased (Bialkova et al., 2013) or the label's surface size should be increased (Peschel et al., 2019).

**Product category and claim types**

Previous studies yielded contradicting results about the influence of the perceived healthiness of product categories on the purchase decision for products labeled with claims. This study tested nutrition, health, and taste claims on three-dimensional packages of orange juice and milk chocolate. It was found that these two categories differed in perceived healthiness for the participants. The purchases of the products were not equal across the different claim types and categories. Orange juices were bought significantly more often with a nutrition claim labeled on the front of the package compared to the labeling with a taste or a health claim. In contrast, milk chocolates were bought significantly more often with a taste claim compared to a nutrition or a health claim.

The results for orange juice support results of previous studies which found that a nutrition claim on a ‘healthy’ food leads to positive evaluations or an increase in purchases (Bialkova et al., 2016; Fenko et al., 2016; Orquin & Scholderer, 2015; Choi et al., 2012). Likewise, the results for milk chocolate support previous studies showing that a taste claim on ‘unhealthy’ food leads to positive evaluations (Fenko et al., 2016; Choi et al., 2012) while a nutrition or a health claim on an ‘unhealthy’ food leads to neutral or even negative effects on preferences, purchase intentions, and actual purchases (Bialkova et al., 2016; Kiesel & Villas-Boas, 2013).

Additionally, the product’s healthiness and nutritional value were rated higher in importance when shopping for orange juice than for milk chocolate. When shopping for milk chocolate, the brand was rated even higher in importance than healthiness and nutritional value. This was confirmed with the purchase behavior (nutrition and health
claims were less preferred to taste claims on milk chocolate) and with the gaze duration (participants looked longer at nutrition tables on orange juice and longer at the brand on milk chocolate). This is in line with studies showing that consumers stated they cared more about taste than health attributes when looking for ‘unhealthy’ food (Chan et al., 2005; Balasubramanian & Cole, 2002). Therefore, the results of the present study provide a further argument in favor of advertising a food product with its strength; a taste claim on chocolate and a nutrition claim on orange juice. In other words, the so-called match-up of claim and product category leads to positive effects.

Health claims were looked at the longest. However, in terms of purchases, orange juices and milk chocolates labeled with a health claim were the least preferred choices. The prominent dislike of health claims on milk chocolates is in line with the lower ratings in trust and belief of the participants compared to orange juice, which in turn is in line with previous studies showing that nutrition and health claims on ‘unhealthy’ food might induce skepticism and distrust (Lalor et al., 2011; Siró et al., 2008). Therefore, this study – as it was a close-to-realistic purchase simulation – adds weight to the argument that health claims might not have a positive effect on evaluations or purchase decisions (Lähteenmäki, 2013; Lähteenmäki et al., 2010).

Conclusions

The results of previous research on nutrition and health claims are contradictory. So far, it has been difficult to draw general conclusions about the impact of claims on consumer purchasing, consumption or even public health. At the same time, stakeholders from the food sector and the policy sector remain very interested in nutrition and health claims (Hieke et al., 2015). The authors of a recent review article suggested testing the claims in more natural situations as the few previous studies conducted in such environments indicated that nutrition and health claims might play a much smaller role than studies conducted in more artificial settings would suggest (Kaur et al., 2017). The present study used a close-to-realistic environment to investigate the effect of different claim types on actual behavior, i.e. the purchase decision and visual attention. The authors recommend continuing research on the effect of nutrition and health claims in close-to-realistic experiments. Based on the results, a recommendation for marketers is to not use health claims because they do not lead to an increase in purchases. The best alternative is the use of nutrition claims which simply state a nutritional benefit of the food. In the case of ‘unhealthy’ products, the use of taste claims is advisable. Previously, it was
pointed out that nutrition and health claims might deceive consumers by outweighing the poor nutritional quality in some food categories (Talati, Pettigrew, Dixon et al., 2016; Orquin & Scholderer, 2015). Whether nutrition or health claims led to a deception about the milk chocolate’s nutritional quality is unknown in this present study. However, the results of the purchase simulation showed that nutrition and health claims were the least preferred claims on milk chocolate.

**Limitations**

Wearing the eye tracking glasses and knowing that one’s eye movements are being observed could potentially influence one’s behavior (Meyerding & Merz, 2018; Graham et al., 2012). Consumers may try to alter their behavior including their natural gaze behavior, e.g. looking less at the graphical elements and longer on more ‘sensible’ elements such as nutrition tables. However, gaze behavior is a subconscious process which is difficult to override (Jacob & Karn, 2003). Avoiding looking at something which attracts attention is rather painful and there is no reason for a participant to execute such a behavior, especially when the participant was unaware of the study’s aim and was instructed to buy the product which they would buy in the supermarket. Since orange juice and milk chocolate were used as two product categories with differing perceived healthiness, it is unknown whether the findings can be generalized to other product categories. The differences in gaze duration obtained on the three different claim types might be due to the differences in word count of the claims and the different complexity in processing them. Furthermore, the experiment was conducted in one German city with a convenience sample unrepresentative of the German population, e.g. participants with university-entrance qualification were overrepresented (61.1% vs. 34.3 %) and participants aged 45 years or older were underrepresented (46.4% vs. 55.9%). Thus, it is unclear whether the findings can be generalized to all German consumers.

**Acknowledgments**

The authors are very grateful to the reviewers who helped to improve the text significantly. Additionally, the authors thank Manika Rödiger for the successful collaboration during the conduction of the study and the preparation of the eye tracking data. The authors also thank Anne Christopherson for proofreading the manuscript. This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors. The authors had no conflicts of interest in writing this paper.
Compliance with ethical standards

The study was in accordance with the ethical standards defined in the 1964 Helsinki Declaration and the study design was approved by the university authorities. Informed consent was obtained from all participants involved in the study. No data was collected that could reveal the identity of the participants.
References


