Food perception and emotion measured over time in-lab and in-home


⁎ Corresponding author.
E-mail address: rene.dewijk@wur.nl (R.A. De Wijk).

Contents lists available at ScienceDirect

Food Quality and Preference
journal homepage: www.elsevier.com/locate/foodqual

Food Quality and Preference 75 (2019) 170–178

ARTICLE INFO

Keywords:
Eating context
Repeated sensory tests
Heart rate
Facial expressions

ABSTRACT

Background: Real-life human eating behaviour does not take place in isolation, rather it happens in a specific physical (e.g. dining room) and social (e.g. family) context. The context in which consumers eat their foods is influential on their hedonic appraisal of the consumed foods (de Graaf et al., 2005; Dijksterhuis, 2016a; Edwards, Meiselman, Edwards, & Lesher, 2003; King, Meiselman, Hottenstein, Work, & Cronk, 2007). Consequently, consumers’ hedonic ratings elicited in a natural consumption context have been observed to differ from those elicited under controlled sensory laboratory conditions (Boutrolle, Delarue, Arranz, Rogeaux, & Köster, 2007; Petit & Siefermann, 2007; Willems, Van Hout, Zijlstra, & Zandstra, 2014). In line with these empirical observations, a close link between the consumption context and how consumers feel has been postulated (Desmet & Schifferstein, 2008), and these feelings are thought to be underlying modulators of food

1. Introduction

Real-life human eating behaviour does not take place in isolation, rather it happens in a specific physical (e.g. dining room) and social (e.g. family) context. The context in which consumers eat their foods is influential on their hedonic appraisal of the consumed foods (de Graaf et al., 2005; Dijksterhuis, 2016a; Edwards, Meiselman, Edwards, & Lesher, 2003; King, Meiselman, Hottenstein, Work, & Cronk, 2007). Consequently, consumers’ hedonic ratings elicited in a natural consumption context have been observed to differ from those elicited under controlled sensory laboratory conditions (Boutrolle, Delarue, Arranz, Rogeaux, & Köster, 2007; Petit & Siefermann, 2007; Willems, Van Hout, Zijlstra, & Zandstra, 2014). In line with these empirical observations, a close link between the consumption context and how consumers feel has been postulated (Desmet & Schifferstein, 2008), and these feelings are thought to be underlying modulators of food

https://doi.org/10.1016/j.foodqual.2019.02.019
Received 20 December 2018; Received in revised form 8 February 2019; Accepted 28 February 2019
Available online 02 March 2019
0950-3293/ © 2019 Elsevier Ltd. All rights reserved.
perception, food liking and overall enjoyment of human eating experiences (Hartwell, Edwards, & Brown, 2013). Consequently, the assessment of foods and drinks in accurate contexts is strongly recommended in order to increase the external validity of the results of these consumer tests (Meiselman, 2013). However, accurate context testing comes at a price, since tests at non-laboratory locations offer typically less control than laboratory tests, which may compromise the validity and reliability of the results.

Typically, home and laboratory testing not only differ in the test location itself but also in a number of other factors that may contribute to differences in test results between home and lab-situations. Bisogni et al. (2007) characterized eating and drinking episodes using eight dimensions: location, food & drink, time, activities, social setting, mental processes, physical condition and recurrence. In most studies the effect of test location itself is confounded with other variables such as social context (eating in isolation in the laboratory versus eating with the family at home) and testing procedures, which makes it difficult to attribute differences between locations in food perception to specific factors. The present study investigates the effect of one specific variable, test location, on hedonic and sensory responses of test foods while other variables are kept constant as much as possible.

Another factor related to test validity is repeated exposure. Research indicated that food acceptance is not static but changes over time and should therefore be assessed in repeated exposures (Köster, 2003; Zandstra & Lion, in press). Researchers have shown an increased liking with repeated consumption, especially relevant for novel, unfamiliar flavours and products (Porchert & Issanchou, 1998), a sustained liking (Kahneinan & Snell, 1992; Levy & Koster, 1999; Willems et al., 2014; Zandstra, Weegels, Van Sprosen, & Klerk, 2004), and a decreased liking (Porchert & Issanchou, 1998; Stubenitsky, Aaron, Catt, & Mela, 1999; Zandstra, De Graaf, & Van Trijp, 2000). Others have related these different effects of repeated exposures to differences in complexity of test foods. Individual consumers may have an optimal level of preferred stimulus complexity. Repeated exposure to levels of complexity that are initially too high may result in increased liking, whereas the opposite may be true for levels of complexity that are initially too low (e.g., Lévy, McRae & Köster, 2006). This and other research suggests at least five to 15 presentations (Dalenberg, Nanneti, Renken, de Wijk, & Ter Horst, 2014; Hetherington, Pirie, & Nabb, 2002; Pliner, 1982; Tijssen, Zandstra, & Jager, 2019; Zandstra et al., 2004, 2018) for a reliable estimate of consumers’ long term dynamics in liking of products.

Food acceptance is typically measured explicitly with self-reported liking, emotions and food perception (Cardello, 1994). Not all food experiences may be accessible for introspection. Alternatives measure food acceptance implicitly using for example facial expressions and responses of the autonomic nervous system (ANS). In food research, facial expressions and ANS responses are typically used as a measure of conscious and unconscious food emotions that play an important role in eating behaviour in everyday life (Canetti, Bachar, & Berry, 2002). Previous studies on foods or food cues demonstrated differential responses of the Autonomic Nervous System (ANS) and/or of facial expressions related to factors such as valence, emotions or degree of satiation (Danner, Haindl, Jochel, & Duerrschmid, 2014; de Wijk, Kooijman, Verhoeven, Holthuysen, & de Graaf, 2012; de Wijk, He, Mensink, Verhoeven, & de Graaf, 2014; He, Boesveldt, de Graaf, & de Wijk, 2016; He, de Wijk, de Graaf, & Boesveldt, 2016; He, Boesveldt, Delpianque, De Graaf & De Wijk, 2017). The speed at which a bite is orally processed and swallowed (consumption duration) was inversely related to the food’s satiating capability. I.e., rapidly consumed foods are less filling and lead to more overconsumption that slowly consumed foods (see de Graaf & Kok, 2010 for a review). It was concluded that implicit responses provide additional information on food preferences when compared to more traditional sensory tests.

Until recently, food studies using this type of implicit measures were restricted to laboratories with well-controlled testing conditions because these measures were technically challenging to perform, were prone to artefacts and all types of interferences, and/or required invasive sensors or elaborate behavioural studies and experimental designs. Recent technical developments using the internet, webcams and non-invasive sensors allow the recording of facial expressions and selected ANS parameters (heart rate) in as well as outside the laboratory (e.g. Urbano, Mahieu, Thomas, Schlich, & Visalli, 2018).

The present study builds on previous studies by investigating the effect of test location (lab versus home) on the evaluation of commercial foods presented repeatedly using a combination of implicit (facial expressions, heart rate and consumption duration) and explicit (sensory profiling and liking) tests.

2. Materials & methods

This study investigated the sensory profile, liking, consumption duration, facial expressions and heart rate for four commercially available foods and one warm-up sample tested repeatedly in home and laboratory test locations.

2.1. Participants

Thirty-two healthy Dutch participants were recruited for this study from the database of Wageningen Food & Biobased Research, social media and printed posters and received a monetary compensation for their participation. In order to be eligible for this study, participants filled out a questionnaire that included questions about drivers of food choice, frequency of meat and meatless consumption, and attitudes towards meat and meat replacers. A selected participant had to meet all of the following criteria: (1) Age between 18 and 65 years, (2) Dutch native language, (3) Healthy (self-reported), (4) BMI between 18.5 and 27 kg/m² (self-reported), (5) Normal ability to taste and smell (self-reported), (6) Being a flexitarian and a light user of soy sauce (≤1x per month and <1x per week), (7) Having given written informed consent, (8) Having access to a Windows laptop with Google Chrome and webcam and willingness to use these in the study. Exclusion criteria were: (1) Being allergic or intolerant to lactose, milk, chocolate, nuts, chicken, meat replacers, tofu, NaCl or soy sauce (self-reported), (2) Being pregnant or lactating (self-reported), and (3) Use medication that may affect the function of taste and smell (self-reported).

The thirty-two recruited participants used foods to substitute meat in their meals. Seventeen (68%) participants ate meat substitutes such as tofu as a substitute, whereas fish (21 (78%)), eggs (18 (72%)), beans (16 (64%)), nuts (13 (52%)), and cheese (12 (48%)) were also reported. Twenty-eight percent of the participants reported that they eat more meat when they are eating with others, whereas 52% reported that this does not influence their meat consumption. Furthermore, participants reported that the aspects taste (4.64 (± 0.64)), healthiness (4.04 (± 0.61)), and price (3.72 (± 0.68)) had the biggest influence on their decision to buy or consume a product (5-point scale). Environment (3.48 (± 0.87)) and convenience (2.92 (± 1.04)) had a smaller impact.

Of the thirty-two recruited participants, only eighteen were included in the study results reported here. Twelve participants were excluded because their data were incomplete, i.e. too many data points were missing of the implicit and/or explicit tests on more than one of the ten test sessions. Most often their incomplete results were caused by technical problems with the internet and/or webcam. Additionally, two participants were excluded based on the poor quality of their video recordings (images were too dark).

2.2. Test foods

This study included samples of four commercially available foods: (1) stir fried chicken breast seasoned with salt (“plain chicken”), (2) stir fried chicken breast seasoned with soy sauce (“soy chicken”), (3) stir fried tofu seasoned with soy sauce (“soy tofu”), and (4) stir fried...
vegetarian chicken breast (Dutch brand: ‘De Vegetarische Slager’) seasoned with soy sauce (“vegi chicken”). As a warm-up sample a piece of minced meat seasoned with salt was used (Dutch brand: Unox), which was a rather different product but with comparable sensory properties particularly concerning its salty taste and meat like origin. The warm-up sample was presented as first sample during each session and corresponding data were not used for further analysis. Participants took two bites (approximately 6 g per bite) per food and evaluated per session a total of 60 g of food, scoring the sensory attributes and liking. All samples were purchased from supermarkets in the Netherlands. The Dutch Food Safety Authority (Nederlandse Voedsel-en Warenautoriteit) ensures that food products sold in the Netherlands are safe for consumption. Samples were prepared in the morning of each test day and stored in the refrigerator at 4 °C according to the hygiene regulations set by the Wageningen University. Samples were served at room temperature in transparent plastic cups labelled with random three-digit codes. Samples were randomized with the software TimeSens® (ChemoSens, Dijon, France) according to a William’s Latin Square experimental design, and the attribute order was kept the same across all participants throughout sensory evaluation.

2.3. General procedure

The study design was a within-subject counterbalanced experiment with five foods (four test foods plus one warm up) that were tested on consecutive weekdays alternating in the participant’s home and in the lab. Half of the participants started testing at home, the other half in the lab (Scheme 1), using their own laptop and webcam.

The presentation order of the test foods was randomized per session and participant. During testing, participants took two bites of each food and scored liking and ten sensory attributes (explicit tests) during which they were video recorded by their own lap tops webcam. These video recordings were used afterwards for the analysis of facial expressions, heart rate and consumption duration (implicit measures). Test sessions lasted about 20 min and were scheduled on a similar time between lunch and dinner. During each session instructions were provided on the computer monitor via internet and the TimeSens® (ChemoSens, Dijon, France) software. The same software was used to present the test items, collect the responses, and to video record the participant’s mimics during sample evaluation. Prior to each evaluation, participants used a video feedback of themselves to verify that the camera position was accurate. Participants were instructed to consume two bites. The first bite was used to evaluate the liking of the food after the food was chewed and swallowed. The second bite was used to evaluate the sensory food properties using a set of 10 sensory attributes. In between the tasting of two samples, participants rinsed their mouth with crackers and water. Each bite was processed at the participant’s own speed. For each bite the sample was put in the mouth while the participant simultaneously pressed the start button on the screen. During the second bite, the intensity of the first attribute shown on the monitor was scored, followed by the next attribute until all attributes were scored. As soon as the test food was swallowed participant pressed a “swallow” button on the monitor which stopped the video recording. A second “stop” button was pressed when the participant was finished scoring. Then, they were instructed to rinse their mouth with water and crackers, and to wait for a period of 3 min before proceeding to the next sample. Participants were instructed to repeat this procedure for all test foods.

2.4. Measurements

2.4.1. Liking

Participants were instructed to rate liking in between bites on a digital visual analogue line scale (VAS) with end anchors “did not like it at all” to “like it very much”.

2.4.2. Sensory attributes

A set of 10 sensory attributes was selected based on sensory characterization of similar foods reported in literature (cf. Imamura, 2016, Dijkstra’s, Luyten, De Wijk, & Mojet, 2007). The attributes contain five taste/flavour attributes (salt, sweet, taste of caramel, taste of bouillon and taste of chicken) and five texture attributes (firm, juicy, speed of disintegration in mouth, oily mouthfeel, and tender). Each attribute was rated on a visual analogue line scale with end anchors “very weak” and “very strong” for the tastes/flavours, “very little” and “very much” for the texture attributes, and “very slow” and “very fast” for the speed attribute. Line scales for all ten attributes were shown simultaneously and in the same order for all test foods and all participants on the monitor.

2.4.3. Facial expressions, heart rate and consumption duration

Video segments of the consumption of each bite were stored together with the participant’s code, the product code, and time and date information. The length of the video segments were used as indicator for consumption duration. Facial expression data were automatically analysed per time frame of 0.04 sec by FaceReader 7.0 (Noldus Information Technology, Wageningen, The Netherlands) in three steps. The face is detected in the first step using the Viola-Jones algorithm (Viola & Jones, 2001). Next, the face is accurately modelled using an algorithmic approach (Den Uyl & Van Kuijlenburg, 2008). Based on the Active Appearance method described by Cootes, Edwards, and Taylor (2001) the model is trained with a database of annotated images that describes over 500 key points in the face and the facial texture of the face. Finally, the actual classification of the facial expressions is based on an artificial neural network trained with 10,000 manually annotated images. The face classification provides the output of seven basic expressions (happy, sad, angry, surprised, scared, disgusted, and contempt), one neutral state on the basis of the Facial Action Coding System developed by Ekman and Friesen (1976), three “affective attitudes” (interest, boredom and confusion), and arousal and valence dimensions based on combinations of facial expressions. FaceReader scores for each emotional expression range from 0 (emotion is not detected) to 1 (maximal detection) and is based on intensity judgments of human experts. FaceReader allows for the simultaneous presence of multiple emotions.

FaceReader was validated by others using the Radboud Faces Database, a standardized test with images of expressions associated with basic emotions. The test persons in the images have been trained to pose a particular emotion and the images have been labelled accordingly by the researchers. Subsequently, the images have been analysed in FaceReader. Accuracy of the assessment of the emotions by FaceReader varied between 84.4% for scared and 95.9% for happy, with an average of 90% Bijnstra & Dotsch, 2011). Yet other validation studies showed superior performance of FaceReader for neutral faces (90% correct recognition for FaceReader versus 59% for humans) (Levinski, 2015). Another study related FaceReader performance to EMG activity of specific facial muscles that have been associated with specific emotions, namely the zygomaticus major or cheek muscle (for happy emotions) and corrugator supercili or brow muscle (for angry emotions). Indeed, FaceReader assessment of happy expressions was significantly correlated (r = 0.72, p < 0.001) with zygomaticus activity, and assessment of angry expressions (r = 0.51, p < 0.05) with corrugator activity.

A more detailed description of the science behind FaceReader can be found at: http://info.noldus.com/free-white-paper-on-facereader-methodology/. Whether all possible emotional expressions can be categorized by these six emotions, affective attitudes and arousal/valence dimensions remains a matter of debate (e.g. Scherer, 2005).

Heart rate was recorded remotely based on Photo-plethysmography, a technique that measures the small changes in color caused by changes in blood volume under the skin epidermis (Wei, Tian, Huang, Wang, & Ebrahim, 2013). Heart rate in combination with other physiological
parameters have previously been related to food preferences and emotions (see de Wijk & Boesveldt, 2016).

2.5. Data analysis

Positions on liking and attribute line scales were converted into scores ranging from zero to ten. Video segments used for analysis of facial expressions and heart rates can be as short as 15 s or as long as 55 s depending on the participant and test food. To verify systematic changes in facial expressions during consumption (e.g., initial expressions of interest may be replaced by expressions of boredom during consumption) differences in durations needed to be normalized first. For normalization, each segment was divided in three tertiles reflecting the start, middle and end of consumption. Per tertile, maximum intensities of expressions and average heart rates were calculated to verify systematic changes during consumption. Statistical analyses were performed with mixed model ANOVAs (using IBM SPSS statistics, version 22) with participant as random factor, and with test food (4), bite (2), test location (2), and replicate per location (5) as fixed factors. Tertile (3) was added as fixed factor for facial expressions and heart rate. Mixed model results as well as the corresponding estimated marginal means will be reported below in the result section.

3. Results

3.1. Liking

Liking scores varied significantly with test food (F(3,654) = 78.3, p < 0.001). Post-hoc tests showed significant differences in liking between all four test foods. Least liked test food was Soya tofu (averaged liking score: 3.5) followed by Vegi Chicken (4.1), soya chicken (5.9) and plain chicken (6.5). Liking scores did not vary with location (F(1,654) = 0.93, n.s.) or with replicate (F(4,654) = 0.51, n.s.).

3.2. Sensory profiles

Scores on each sensory attribute varied with test food (p < 0.001, see Fig. 1). The largest differences were found for the texture attributes whereby tofu was judged to be less firm, more tender and faster desintegrating in the mouth than the other test foods. With regard to the flavor attributes, not suppringly chicken flavor was more intense in both chicken foods than in the other foods. Saltiness was more intense in the foods with soy sauce. The scoring on sensory attributes did not vary with test location, except for oily (F(1,654) = 5.2, p = 0.02) in that oily was perceived less intense in sensory lab compared to home.

Repeated exposure (from session 1 to 5) resulted in reduced ratings of bouillon (F(4,654) = 5.7, p < 0.001) and salt (F(4,654) = 11.8, p < 0.001), and increased ratings of tenderness ((F(4,654) = 4.4, p = 0.002)) (see Fig. 2).

3.3. Consumption duration

Consumption duration varied with test sample (F(3,1243) = 20.3, p < 0.001), with fastest consumption for the soya tofu (25.9 sec, p < 0.001) and slower and similar durations for the other three test foods. Consumption at home was faster than consumption in the lab (29.0 vs. 30.1 sec. F(1,1245) = 4.2, p = 0.04), and consumption became faster with replication (average duration 33.3 s for rep 1 to 26.7 s for rep 5) (F(4,1244) = 15.4, p < 0.001). Consumption of bite 1 (liking) was faster than of bite 2 (sensory profile) (17.8 vs 41.3 s, F(1,1244) = 1796, p < 0.001) (see Table 1). Correlational analyses showed positive associations between durations and specific facial expressions (sad, angry, surprise, scared, r > 0.34, p < = 0.05, n = 50), specific affective states (arousal, boredom and interest, r > 0.28, p < = 0.05, n = 50), and with one sensory attribute (firmness, r = 0.28, p < = 0.05, n = 50). Duration was not associated with liking (r = 0.21, n.s., n = 50).

3.4. Facial expressions

The results of the statistical analyses of the facial expressions, dimensions and affective attitudes are summarized in Table 2. Facial expressions that varied with test food are sad, surprise, and scared. The valence dimension varied as well with test food (Fig. 3). In general, these facial expressions were least intense for tofu. Facial expressions, dimensions and affective states varied with test location (Fig. 4). At home, higher intensities were found for happy, disgust, contempt and boredom, and also for valence, the combination of positive and negative facial expressions. The other expressions and affective attitudes are more intense in the lab (Table 2). During consumption, i.e. across tertiles, negative expressions become stronger (sad, scared, contempt and boredom). Arousal and valence decrease (Fig. 5). Across exposure, expressions and affective states, except sad, boredom and confusion, become less intense.

3.5. Heart rate

Heart rate varies between samples (F(3,2955) = 24.1, p < 0.001) with highest heart rate for plain chicken (74.6 ± 0.9 beats per minute or BPM), lower heart rate for soy and vegi chicken (71.4 ± 0.9 BPM;
71.9 ± 0.9 BPM), and lowest heart rates for soya tofu (69.9 ± 0.9 BPM, p < 0.05). Heart rates during home testing were higher than during lab testing (72.6 ± 0.9 vs 71.3 ± 0.9 BPM, F(1,2967) = 10.5, p = 0.001). During consumption, heart rates gradually decreased from 73.1 ± 0.9 to 71.7 ± 0.9 and 70.8 ± 0.9 BPM (tertile effect: F(2,2954) = 11.9, p < 0.001). Heart rate varies with replicate (F(4,2957) = 2.3, p = 0.05). Heart rates during the second replicate (70.9 ± 0.9 BPM) were significantly lower than those during the first (72.2 ± 0.9 BPM) third (72.4 ± 0.9 BPM and fifth (72.5 ± 1.0 BPM) replicate (p < 0.03).

Correlational analysis showed strong negative associations between heart rate and the sensory texture attributes juiciness, speed and tenderness (r < 0.51, p < = 0.01, n = 50), positive association with the texture and flavor attributes firmness and chicken flavor (r > = 0.54, p < = 0.01, n = 50), and moderate associations with specific facial expressions (happy, disgust, valence, r > = 0.28, p > = 0.05, n = 50) and affective states (interest, r = −0.30, p = 0.05, n = 50). Heart rate was not associated with duration (r = −0.04, n.s., n = 50) or liking (r = 0.26, n.s., n = 50).

Table 1
Consumption durations in seconds (± S.D.) per product, test location, bite and replicate.

<table>
<thead>
<tr>
<th>Location</th>
<th>Bite</th>
<th>Replicate</th>
<th>Plain chicken</th>
<th>Soya chicken</th>
<th>Soya tofu</th>
<th>Vegi chicken</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home 1</td>
<td>1</td>
<td>1</td>
<td>21.12 ± 3.6</td>
<td>23.85 ± 3.6</td>
<td>17.13 ± 3.6</td>
<td>22.82 ± 3.6</td>
</tr>
<tr>
<td>Home 1</td>
<td>2</td>
<td>1</td>
<td>17.63 ± 3.5</td>
<td>20.35 ± 3.5</td>
<td>13.74 ± 3.5</td>
<td>18.85 ± 3.5</td>
</tr>
<tr>
<td>Home 1</td>
<td>3</td>
<td>1</td>
<td>18.14 ± 3.6</td>
<td>18.19 ± 3.5</td>
<td>11.24 ± 3.5</td>
<td>17.84 ± 3.6</td>
</tr>
<tr>
<td>Home 1</td>
<td>4</td>
<td>1</td>
<td>19.13 ± 3.5</td>
<td>18.13 ± 3.5</td>
<td>12.14 ± 3.6</td>
<td>22.30 ± 3.5</td>
</tr>
<tr>
<td>Home 1</td>
<td>5</td>
<td>1</td>
<td>15.21 ± 3.7</td>
<td>15.32 ± 3.6</td>
<td>10.57 ± 3.6</td>
<td>17.76 ± 3.6</td>
</tr>
<tr>
<td>Home 2</td>
<td>1</td>
<td>2</td>
<td>45.88 ± 3.8</td>
<td>43.47 ± 3.7</td>
<td>40.11 ± 3.6</td>
<td>45.97 ± 3.6</td>
</tr>
<tr>
<td>Home 2</td>
<td>2</td>
<td>2</td>
<td>43.16 ± 3.7</td>
<td>42.38 ± 3.6</td>
<td>40.60 ± 3.6</td>
<td>45.54 ± 3.6</td>
</tr>
<tr>
<td>Home 2</td>
<td>3</td>
<td>2</td>
<td>35.17 ± 3.7</td>
<td>40.30 ± 3.6</td>
<td>39.30 ± 3.6</td>
<td>39.80 ± 3.6</td>
</tr>
<tr>
<td>Home 2</td>
<td>4</td>
<td>2</td>
<td>40.64 ± 3.7</td>
<td>37.63 ± 3.6</td>
<td>33.87 ± 3.6</td>
<td>40.46 ± 3.6</td>
</tr>
<tr>
<td>Home 2</td>
<td>5</td>
<td>2</td>
<td>35.59 ± 3.8</td>
<td>40.76 ± 3.7</td>
<td>37.88 ± 3.6</td>
<td>40.16 ± 3.7</td>
</tr>
<tr>
<td>Lab 1</td>
<td>1</td>
<td>1</td>
<td>22.58 ± 3.5</td>
<td>25.00 ± 3.5</td>
<td>15.94 ± 3.5</td>
<td>21.21 ± 3.5</td>
</tr>
<tr>
<td>Lab 1</td>
<td>2</td>
<td>1</td>
<td>18.63 ± 3.5</td>
<td>19.47 ± 3.5</td>
<td>14.26 ± 3.5</td>
<td>21.68 ± 3.5</td>
</tr>
<tr>
<td>Lab 1</td>
<td>3</td>
<td>1</td>
<td>18.91 ± 3.6</td>
<td>20.14 ± 3.6</td>
<td>13.43 ± 3.6</td>
<td>18.76 ± 3.6</td>
</tr>
<tr>
<td>Lab 1</td>
<td>4</td>
<td>1</td>
<td>20.08 ± 3.6</td>
<td>17.85 ± 3.6</td>
<td>12.55 ± 3.6</td>
<td>21.43 ± 3.6</td>
</tr>
<tr>
<td>Lab 1</td>
<td>5</td>
<td>1</td>
<td>14.69 ± 3.6</td>
<td>14.38 ± 3.6</td>
<td>12.63 ± 3.6</td>
<td>16.63 ± 3.6</td>
</tr>
<tr>
<td>Lab 2</td>
<td>1</td>
<td>2</td>
<td>54.11 ± 3.5</td>
<td>44.28 ± 3.5</td>
<td>43.05 ± 3.6</td>
<td>45.88 ± 3.6</td>
</tr>
<tr>
<td>Lab 2</td>
<td>2</td>
<td>2</td>
<td>46.49 ± 3.5</td>
<td>44.37 ± 3.5</td>
<td>40.11 ± 3.5</td>
<td>42.11 ± 3.6</td>
</tr>
<tr>
<td>Lab 2</td>
<td>3</td>
<td>2</td>
<td>41.69 ± 3.5</td>
<td>40.84 ± 3.6</td>
<td>36.85 ± 3.6</td>
<td>41.30 ± 3.5</td>
</tr>
<tr>
<td>Lab 2</td>
<td>4</td>
<td>2</td>
<td>42.79 ± 3.6</td>
<td>39.74 ± 3.6</td>
<td>37.09 ± 3.6</td>
<td>45.57 ± 3.5</td>
</tr>
<tr>
<td>Lab 2</td>
<td>5</td>
<td>2</td>
<td>38.20 ± 3.7</td>
<td>43.29 ± 3.7</td>
<td>35.34 ± 3.7</td>
<td>41.80 ± 3.7</td>
</tr>
</tbody>
</table>

4. Discussion

Test foods were tested repeatedly at home and in the laboratory using a set of explicit (liking and sensory attributes) and implicit (facial expressions, heart rate and consumption duration) measures. In general, implicit measures showed effects of test food, test location, and exposure (replication). Explicit measures also showed effects of test food (on liking and sensory attributes) An additional effect of exposure on specific attributes (which) appeared. None of the explicit tests showed effects of test location. Hence, in this study explicit tests appear to be less sensitive to variations in the experimental context (e.g. in test location) than implicit tests. Traditional product testing typically involves explicit tests. So these tests run the risk of missing effects of the experimental context.

The fact that most new food products disappear quickly from the market despite exhaustive explicit testing (cf. Dijksterhuis, 2016b) may demonstrate the limited relevance of explicit tests for long term consumer acceptance. The higher sensitivity of implicit measures found in this study is encouraging but future studies should verify the relevance of this higher sensitivity for consumer acceptance, for example by relating implicit and explicit test measures to longer term choice and sales.

Fig. 2. Sensory attributes (bite 2) from session 1 to 5 averaged across participants, test foods and locations. Significant differences between replicates are indicated with “s”, non-significant differences with “ns”.

Table 1
Consumption durations in seconds (± S.D.) per product, test location, bite and replicate.
data.

The liking results in this study failed to show systematic effects of repeated exposure and consumption location. Previous studies have demonstrated systematic changes in liking during repeated exposures and these changes have been related for example to the food’s perceived complexity. The relative small number of participants in this study, as well as the fact the test foods were not selected based on their complexity, does not allow verification of this relationship in the data from this study. Possible explanations for the lack of location effects include low statistical power as a result of the small number of participants included in this study or the used within-subject study design where the same participant is tested repeatedly alternating at home and in the lab with the same test products. Memories of previous encounters with the same test food may have facilitated the use of similar liking scores in new encounters. As a result liking scores would show an artificially high degree of stability over occurrences. In contrast, other measures such as

Table 2
summary of fixed effects from mixed model ANOVA of facial expressions, dimensions and affective attitudes.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Neutral</th>
<th>Happy</th>
<th>Sad</th>
<th>Angry</th>
<th>Surprised</th>
<th>Scared</th>
<th>Disgusted</th>
<th>Contempt</th>
<th>Valence</th>
<th>Arousal</th>
<th>Interest</th>
<th>Boredom</th>
<th>Confusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1,363</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Bite</td>
<td>1,363</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>SampleName</td>
<td>3,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>Tertile</td>
<td>2,363</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Testlocation</td>
<td>1,363</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>RepPerLocation</td>
<td>4,363</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Bite * SampleName</td>
<td>2,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>Bite * Tertile</td>
<td>1,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>Bite * Testlocation</td>
<td>1,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>Bite * RepPerLocation</td>
<td>4,363</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>SampleName * Tertile</td>
<td>6,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>SampleName * Testlocation</td>
<td>3,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>SampleName * RepPerLocation</td>
<td>12,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>Tertile * Testlocation</td>
<td>2,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>Tertile * RepPerLocation</td>
<td>8,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>Testlocation * RepPerLocation</td>
<td>4,363</td>
<td>0.386</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bite * SampleName * Testlocation</td>
<td>3,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>Bite * SampleName * RepPerLocation</td>
<td>12,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>Bite * Tertile * Testlocation</td>
<td>2,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>Bite * Tertile * RepPerLocation</td>
<td>8,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>Bite * Testlocation * RepPerLocation</td>
<td>4,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>SampleName * Tertile * Testlocation</td>
<td>6,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>SampleName * Tertile * RepPerLocation</td>
<td>24,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>SampleName * Testlocation * RepPerLocation</td>
<td>12,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>SampleName * Tertile * Testlocation</td>
<td>12,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>SampleName * Tertile * RepPerLocation</td>
<td>12,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>SampleName * Testlocation * RepPerLocation</td>
<td>12,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>SampleName * Tertile * Testlocation</td>
<td>12,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>SampleName * Tertile * RepPerLocation</td>
<td>12,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
<tr>
<td>SampleName * Testlocation * RepPerLocation</td>
<td>12,363</td>
<td>N.S.</td>
<td>N.S.</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
<td>N.S</td>
</tr>
</tbody>
</table>

N.S.: not significant.
* 0.01 < p < 0.05.
** 0.001 < p < 0.01.
*** p < 0.001.

Fig. 3. Facial expressions, dimensions and affective attitudes per test food averaged across participants, sessions and locations. Significant differences between foods are indicated by “s”, non-significant differences by “ns”.
sensory profiles, and especially behaviour (duration) and facial expressions were much more responsive to the effects of location and exposure, providing support for the possibility that the stability of liking scores is at least partly driven by memory. Alternatively, the difference between home and lab conditions in the present study may have been smaller than the differences in other studies. In the present study, participants consumed the same quantity of the same test foods at similar times in isolation at home or in the lab, responding to the same questions using the same laptop and webcam. Future studies should systematically investigate the contributions of other dimensions on food appreciation. In contrast to the explicit measures, the implicit measures used in this study (consumption duration, facial expressions and heart rate) showed systematic effects of not only sample but also location and exposure. The fact that these measures show effects that are not reflected in liking scores could imply that the processes reflected by the implicit measures may not be the same as those responsible for explicit liking and sensory attribute scores. They may therefore be valuable additions in the toolbox to assess consumer responses and food acceptance, although it remains to be shown that these implicit measures can meaningfully be related to (long-term) acceptance. Alternatively, the discrepancy between liking and implicit measures may provide evidence for the hypothesis stated earlier, namely that repeated liking scores reflect (at least partly) memories of previous encounters with the test samples, rather than actual experiences at the moment of testing.

Heart rate was especially responsive to variations in sample, location and replication. In addition, heart rate drops gradually during consumption. Heart rate variations could reflect differences in affective responses. Previous research showed reduced heart rates for well-liked stimuli such as pleasant aromas (Alaoui-Ismaili et al., 1997a, 1997b; He, Boesveldt, de Graaf, & de Wijk, 2014), even though this relationship seems to be less clear for food (consumption) (Brouwer, Hogervorst, Grootjen, Van Erp, & Zandstra, 2017; de Wijk & Boesveldt, 2016). Moreover, in this study the lowest heart rate was observed for the least-liked (soya tofu) rather than the most-liked test food, which prompts an alternative explanation. Probably, the heart rates measured in the present study may reflect primarily physical muscle (chewing) activity. Foods that require more and intense chewing should result in higher heart rates than other foods. Physical properties such as toughness were not measured directly but rather indirectly via sensory attributes such as firmness. The strong correlations between heart rate and perceived firmness ($r = 0.59$), speed of disintegration ($r = -0.54$) and

---

**Fig. 4.** Facial expressions, dimensions and affective attitudes per test location averaged across test foods and participants.

**Fig. 5.** Changes in facial expressions, dimensions and affective attitudes during consumption (tertiles 1–3) averaged across test foods and participants.
tenderness ($r = -0.51$) support the hypothesis that in this study heart rate may mainly reflect chewing muscle activity. In line with this activity hypothesis is the observation that heart rate gradually decreases during consumption: during consumption, foods lose their consistency primarily due to mechanical degradation. Degraded foods require less muscle activity resulting in lower heart rates. The fact that eating at home was associated with higher heart rates than eating in the lab suggests that foods are consumed at home with more effort than in the lab. Possibly, foods consumed at home are chewed more vigorously than foods in the lab. Interestingly, more vigorous chewing was not related to faster eating (no association between heart rate and duration), and did not affect sensory attributes.

Even though the differences between the home and lab testing conditions were relatively small, as was discussed earlier, these differences were large enough to show significant changes in eating behaviour. Test foods consumed at home were consumed at a faster rate than the same foods consumed in the laboratory. Not only eating behaviour varied with location but also the associated facial expressions and heart rate suggest that foods are chewed faster with greater chewing intensity at home compared to in the lab.

Previous research has demonstrated that variations in speed of eating can affect the perception of the food but this typically concerned variations between consumer segments that each display their own idiosyncratic way of eating (Devezeaux de Lavergne, Derks, Ketel, de Wijk, & Stieger, 2015). Research where the effect of variations in speed eating on perception is investigated in the same consumers, similar to the present study, is very scarce and shows perceptual changes when consumers change their eating behaviour (de Wijk, Engelen, & Prinz, 2003; de Wijk, Janssen, & Prinz, 2011). The fact that in this study changes in speed of eating are not always accompanied by perceptual changes provides additional evidence for the involvement of memory processes in perception as stated previously.

5. Conclusion

Liking and attribute ratings of foods tested explicitly and repeatedly at home and in the lab showed differences between foods but not between test locations. In addition, specific sensory attribute ratings showed systematic effects over repeated exposure. Interestingly, implicit tests (facial expressions, heart rate and consumption duration) showed systematic differences between test foods, over repeated exposure and between test locations. Future studies should investigate the relevance of the increased sensitivity of implicit tests for long-term consumer acceptance, for example by relating implicit and explicit test results to long term food acceptance, repeated food choice or actual product sales.

Acknowledgements

This work was supported by a grant from the Dutch Top Consortium for Knowledge and Innovation (TKI) Agri&Food together with Unilever R&D Vlaardingen, The Netherlands, Kikkoman Europe R&D Laboratory and Noldus Information Technology (TKI-AF-17005).

References


