Does the face show what the mind tells? A comparison between dynamic emotions obtained from facial expressions and Temporal Dominance of Emotions (TDE)

Roelien van Bommel[a,b,*], Markus Stieger[a,b], Michel Visalli[c], Rene de Wijk[d], Gerry Jager[a,b]

[a] TiFN, P.O. Box 557, 6700 AN Wageningen, The Netherlands
[b] Wageningen University, Division of Human Nutrition and Health, P.O. Box 17, 6700 AA Wageningen, The Netherlands
[c] INRA, UMR 1324 Centre des Sciences du Goût et de l'Alimentation, INRA, 9e bd Jeanne d'Arc, F-21000 Dijon, France
[d] Food & Biobased Research, Wageningen University, Bornse Weilanden 9, 6708WG Wageningen, The Netherlands

ARTICLE INFO

ABSTRACT

Measuring food-evoked emotions dynamically during consumption can be done using explicit self-report methods such as Temporal Dominance of Emotions (TDE), and implicit methods such as recording facial expressions. It is not known whether or how dynamic explicit and implicit emotion measures correspond. This study investigated how explicit self-reported food-evoked emotions evaluated with TDE are related to implicit food-evoked emotions determined from facial expressions. Fifty-six participants evaluated six yogurts with granola pieces varying in size, hardness and concentration, using multiple bite assessment employing TDE for the first, third and fifth bite of consumption. Consumers were video recorded during each bite of consumption and facial expressions were analysed using FaceReader™. *Happy, interested, disgusted* and *bored* were similar descriptors measured explicitly and implicitly. Little overlap was observed regarding the type of emotion characterization by FaceReader™ and TDE. Products were mainly discriminated along the valence dimension (positive – negative), and directly reflected product discrimination in terms of liking. FaceReader™ further differentiated the least liked products from each other on arousal and negative facial expressions. Our results indicated little dynamics in food-evoked emotions within and between bites. Facial expressions seemed more dynamic within bites, while explicit food-evoked emotion responses seemed more dynamic between bites. We conclude that FaceReader™ intensities of emotions and dominance durations observed in TDE are not directly comparable and show little overlap. Moreover, food-evoked emotion responses were fairly stable from first to last bite and only very limited changes were observed using implicit and explicit emotions measures.

1. Introduction

Sensory perceptions of foods and beverages change dynamically during consumption due to mastication and salivation (Castura, Antúnez, Giménez, & Ares, 2016; Delarue & Blumenthal, 2015; Pineau et al., 2009). Consequently, changes in appraisal of these dynamic sensory perceptions might lead to an unfold of different food-evoked emotions during consumption. The Component Process Model (CPM) by Scherer (2005, 2009) describes emotions as dynamic events that change upon the cognitive appraisal of a stimulus (e.g. food) (Fig. 1). The CMP defines emotions as dynamic episodes, with an onset (event, stimulus) followed by a complex process of continuous changes both centrally in the brain, and peripherally via the co-occurring bodily symptoms and expressions (e.g. heart rate, blood pressure, and facial

and vocal expressions), and eventually the subjective, conscious experience, the feeling one becomes aware of (Jager, 2016; Scherer, 2005, 2009).

It is suggested that self-report measures only reveal the emotion one becomes aware of, whereas parts of the complex emotion process in other subsystems remain hidden (Kahneman, 2003; Köster, 2003; Köster & Mojet, 2015; Scherer, 2005, 2009). More implicit measures, such as facial expressions might provide additional information on fast changing emotions during food consumption. Few studies compared the performance of facial expressions and self-reported food-evoked emotion measurements (He, Boesveldt, de Graaf, & de Wijk, 2016; Leitch, Duncan, O'keefe, Rudd, & Gallagher, 2015). Leitch et al. (2015) compared product profiles of natural and artificial sweeteners in tea obtained with a self-reported emotion questionnaire (Check-All-That-
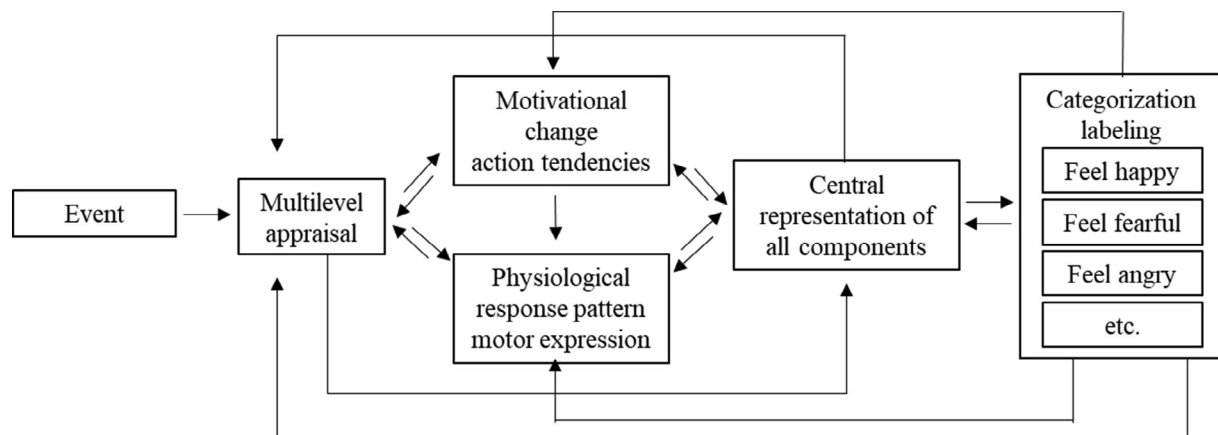
Fig. 1. Component Process Model (CMP). .
Source: Scherer (2009)

Apply) and facial expressions (FaceReader™, version 5.0). They observed product differentiation using the emotion questionnaire, but they did not find significant differences in facial expression profiles between products (Leitch et al., 2015). He et al. (2016) compared an explicit non-verbal emotion method (PrEmo®) with facial expressions (FaceReader™, version 4.0). They concluded that the self-reported food-evoked emotions are relatively unidimensional, whereas facial expressions report multidimensional aspects such as intensity and the sequential unfolding of emotions during food consumption.

An explicit method that allows consumers to self-report dynamic changes in emotion perception during tasting is the Temporal Dominance of Emotions (TDE) methodology (Jager et al., 2014; Mahieu, Visalli, Schlich, & Thomas, 2019). TDE originates from the Temporal Dominance of Sensations (TDS) technique, and is based on the concept of dominance (e.g. defined as the emotion catching most of the attention at each time) (Jager et al., 2014; Pineau et al., 2009). TDE might provide a better dynamic understanding of a consumer's subjective product experience because it allows the sequential evaluation of the perceived food-evoked emotions that dominate during consumption.

Different components of the emotion process are complementary, and linking implicit to explicit emotion measurements over time will generate novel insights on how to interpret consumers' affective responses in relation to food and eating behaviour. Previous findings on dynamic changes of sensory perceptions using multiple bite assessments, indicate that different food components dynamically interact with one another during consumption and evoke a perceptual change in sensory characteristics from bite to bite (van Bommel, Stieger, Boelee, Schlich, & Jager, 2019). Hence, exposure to multiple bite intakes impacts the temporal dynamics of sensory perceptions, and consequently, may elicit a change in hedonic and emotion evaluations, both within- and between bites. To investigate this, we recorded facial expressions during the subjective evaluation of six yogurts with added granola varying in hardness, size and concentration employing TDE and TDS using a five bite evaluation approach. Sensory profiles of the yogurt with added granola, presented in a separate paper, revealed product differentiation between samples on hardness of the granola particle and on the concentration of granola added to the yogurt (van Bommel et al., 2019). The different sensory characteristics of these yogurts with added granola would lead to differential emotion profiles. This study aims to compare dynamic changes in emotion profiles and product discrimination employing implicit (facial expressions) and explicit (TDE) emotion measures. Although the type of information obtained with monitoring facial expressions and TDE is very different and, therefore, not directly comparable, we hypothesized a certain extent of correspondence between both emotion components. We hypothesized that

results (i.e. dynamic changes) measured by both methods correspond at the level of a two-dimensional framework of valence (positive – negative) and arousal (high activation – low activation) within and between bites (Russell, 1980).

2. Materials and methods

As part of a larger study, participants completed two separate test sessions; one for the sensory and hedonic evaluations employing Temporal Dominance of Sensations (TDS) and alternated-Temporal Drivers of Liking (a-TDL), and a second session for emotion evaluations employing Temporal Dominance of Emotions (TDE). Simultaneously with these sessions participants were video recorded in order to monitor facial expressions using FaceReader™ (version 7.0, Noldus Information Technology, Wageningen, The Netherlands). The data and findings on sensory perceptions and drivers of liking (TDS and alternated-TDL) are outside the scope of the current paper and have been reported elsewhere (van Bommel et al., 2019). This paper focuses on food-evoked emotion evaluations employing TDE and FaceReader™ (version 7.0). All data were collected at Wageningen University (The Netherlands). The experimental protocol was submitted to and exempted from ethical approval by the medical ethics committee of Wageningen University.

2.1. Participants

Seventy-six healthy Dutch participants, between 18 and 65 years old, participated in this study. After data collection, participants with more than 5% missing data frames were removed from data analysis. Consequently, twenty participants were excluded from data analysis resulting in a total of fifty-six participants (17 male, 39 female, mean age 27.7 ± SD 11.9 years, mean BMI 22.1 ± SD 2.1 kg/m^2) included in the data analysis of this study. Incomplete FaceReader™ data frames were caused by a loss of eye contact with the camera; inappropriate lighting that caused shadows in the face which made it impossible for FaceReader™ to quantify the facial expression; people wearing glasses; and, people with facial hair, such as beards and moustaches. Participants were recruited from a database with volunteers to participate in research of the Division of Human Nutrition of Wageningen University, the Netherlands. All participants were consumers of yogurt, without allergies or intolerances for lactose, gluten, milk or nuts and with normal abilities to taste and smell (self-reported). Participants received a monetary incentive for their participation, and gave written informed consent before the start of the study.

Table 1
Product specifications.

Product	Granulation (mm)	Hardness	Concentration		kcal
			Yogurt (g)	Granola (g)	
Hard:Large:10%	23	Hard	54	6	57
Hard:Small:10%	10	Hard	54	6	57
Soft:Large:10%	23	Soft	54	6	57
Soft:Small:10%	10	Soft	54	6	57
Hard:Small:20%	10	Hard	48	12	84
Hard:Small:3%	10	Hard	58	2	38

Granulation: size of the breaking grids that were used to define particle sizes.

2.2. Products

Commercially available yogurt (Optimel Greek Style, Friesland Campina, The Netherlands) with commercially available granola (Crunchy Hazelnut Granola, Biofamilia, Switzerland) were used. Composite food (i.e. combination of two or more foods) were chosen because of their increased sensory complexity as the sensory characteristics of one food product includes the sensory perceptions of the other food. Product characteristics are specified in Table 1. Yogurt with granola samples differed in hardness (hard vs. soft), particle size $(9.5 \pm 0.22$ mm vs. 19.7 ± 0.24 mm) and concentration (3%, 10% and 20%) added to the yogurt. For more details on the product characteristics, see van Bommel et al. (2019). Participants received a total of 60 g per yogurt-granola combination, presented in white plastic cups coded with 3-digits. A warm-up sample, consisting of 54 g yogurt with 3 g of small granola and 3 g of large granola, was included to familiarize participants with the study procedures.

2.3. Attribute selection

FaceReader™ is able to detect 6 basic emotions (*angry, contempt, disgusted, happy, scared* and *surprised*), a neutral state (*neutral*) and 3 affective attitudes (*interest, bored* and *confused)*. To allow comparison with facial expression analysis by FaceReader™ the emotions *bored, disgusted, interested* and *happy* were included in the TDE evaluations. Twenty emotion attributes were preselected based on literature (Gutjar et al., 2015; King & Meiselman, 2010; Schouteten, De Steur, Sas, De Bourdeaudhuij, & Gellynck, 2017). A Check-All-That-Apply was performed by 10 consumers (not participating in real experiment). The 6 most frequently cited emotion attributes were used in this study together with the four preselected emotion terms mentioned above. Table 2 shows the emotion attributes with descriptions as provided to the participants during TDE instructions.

2.4. Procedure

Participants completed two test sessions for the emotion evaluations. Each session, participants evaluated one warm-up sample and

three test samples. The total amount of product evaluated per session was 240 g, which approximately corresponds to the amount of a full portion. Sessions lasted about 45 min and were scheduled on separate days between 08.00 and 10.00 h. Participants conducted the emotion evaluations on the same time of day. Sessions took place in sensory booths (Restaurant of the Future, Wageningen, The Netherlands). Sensory booths were designed according to ISO 8589 standards (ISO, 2007), and tests were conducted under artificial daylight and temperature control (20–22 °C). One day before each session participants received the attribute list with definitions by email to familiarize themselves with the terminology. A live demonstration of the study procedures was given at the start of the first session. Participants were instructed to consume the whole sample (60 g) in five bites, and to always consume yogurt and granola within one bite. All bites were video recorded, and participants performed TDE for the first, third and fifth bite using TimeSens software (version 1.1.601.0, ChemoSens, Dijon, France). During the second and fourth bite ('no task') participants just ate the bite without performing TDE or liking ratings. When perception ended, participants had to click the stop button, allowing time and video recording to stop. After the first, third and fifth bite participants were instructed to rate liking on a continuous scale with end anchors 'dislike extremely' and 'like extremely'. A 3 min neutralisation period was included between samples where participants ate a piece of cracker and rinsed their mouth with water.

2.5. Dynamic emotion measurements

2.5.1. Temporal dominance of emotions

Participants were instructed to put a full spoon with yogurt and granola into their mouth and simultaneously click the start button, allowing time recording to start. Then, they had to select the dominant attribute (e.g. the attribute that catches most of their attention), and dominance recording of that attribute started from then and remained selected until a new dominant attribute was selected. When perception ended, participants had to click the stop button, allowing time recording to stop (Pineau et al., 2009). Participants could select as many dominant attributes as they liked, using the same attributes several times or never select an attribute during the consumption time.

2.5.2. FaceReader™

Participants were video recorded using a Logitech C270 webcam with a resolution of 720p mounted on top of the computer screen. FaceReader™ (version 7.0, Noldus Information Technology, Wageningen, The Netherlands) was used to automatically classify facial expressions from the video recordings at a time frame of 0.02 s. Upon facial recognition, an artificial 3D face model is obtained based on the Active Appearance Modelling (AAM) (Cootes, Edwards, & Taylor, 2001) using 500 key points in the face. For each data frame, facial expressions are classified based on a database of 10.000 facial expression images that were manually classified by trained experts. Deep Face classification method was used to allow facial expression recognition when their

Table 2
Emotion attributes and descriptions/examples.

Emotion	Dimension	Description	Example
Aggressive	Arousal	Destructive behaviour mostly caused by frustration	Losing a game makes me feel aggressive
Bored	Arousal	Finding something uninteresting.	Doing the same thing every day makes me feel bored
Calm	Arousal	A state of freedom from excitement or disturbance	Yoga and meditation make me feel calm
Disgusted	Valence	A strong aversion to something or someone.	Closing a full garbage bag makes me feel disgusted
Energetic	Arousal	Having or showing energy	I feel energetic after a good night sleep
Enthusiastic	Arousal	Excited or exuberant feeling	I feel enthusiastic when I score a goal
Good	Valence	Pleasant or good feeling	Helping someone makes me feel good
Happy	Valence	To be pleased or glad	After passing my exam I felt very happy
Interested	Valence	To arouse or hold an interest in someone or something	The claim on the package label made me interested in the product
Whole	Valence	Seemingly complete or total	Being with family makes me feel whole

eyes were still identifiable but when the lower part of the face is hidden (e.g. when they cover the mouth with a spoon). Detailed information on how facial expressions are identified with FaceReader™ is described in the FaceReader™ Methodology Note by Loijens and Krips (https://info. noldus.com/free-white-paper-on-FaceReader-methodology). Emotions and attitudes are given a score between 0 (absent) and 1 (fully present) depending of the intensity of the facial expression. Furthermore, FaceReader™ calculates *valence* (i.e. positive or negative emotion state) and *arousal* (i.e. level of activation). Valence is scored between -1 (negative emotions) and 1 (positive emotions), and arousal is scored between 0 (not active) and 1 (active).

2.6. Data analysis

Statistical analysis was performed using R (R version 3.4.2, RStudio team, 2016). Results were considered significant at $p < 0.05$, unless stated otherwise.

Dominance durations, maximum intensities of facial expressions and liking scores were checked for first order effect across serving positions. No significant order effects of serving position was observed (data not reported). Therefore, product order was no longer included in the mixed model ANOVA for dominance durations, maximum facial expressions and liking.

2.6.1. Temporal dominance of emotions

TDE bandplots were plotted using TimeSens software (version 1.1.601.0, ChemoSens, Dijon, France). Bandplots represent the sequence and duration of significant dominant attributes as time-bands (Galmarini, Visalli, & Schlich, 2017), and were computed by product for the first, third and fifth bite. Coloured rectangles represent the dominant attributes and are stacked at each moment, displaying multiple dominances (without taking into account dominance rates at a given time point). The total height of the band is a constant and the number of colours at each moment depends on the number of significantly dominant attributes at the same time, providing a characteristic 'patchwork' effect. TDE bandplots were visually inspected to identify differences and similarities in dominance sequences between products.

The first, third and fifth bites were divided into three periods (i.e. beginning, middle and end of a bite). Mean dominance durations and standard errors of the mean were calculated per tertile, bite and product for each emotion attribute. A mixed model ANOVA was performed with product, bite and tertile as fixed factors and subject and its interaction effects with all fixed factors as random effects. Tukey HSD pairwise comparison was performed upon significance of the ANOVA.

2.6.2. Facial expressions

Facial expression data was quantified using FaceReader™ (version 7.0) at a frequency of 5 Hz (i.e. 5 data frames per second) using the 'general face. Individual calibration was not used since the study followed a within-subject design. All subjects evaluated all samples in all conditions. This allows to directly quantify changes in facial expressions caused by samples in all conditions without calibration. Calibration of individual facial expression responses to a neutral stimulus to correct for potential biases in an individual's facial response were therefore not employed. Data was standardized by dividing each bite into three periods (i.e. beginning, middle and end of a bite). Maximum intensities and standard errors of the mean were calculated per tertile, bite and product for each facial expression. A mixed model ANOVA was performed with product, bite and tertile as fixed factors and subject and its interaction effects with all fixed factors as random effects. Upon significance of the ANOVA, a Tukey HSD pairwise comparison was performed.

2.6.3. Comparison between Temporal dominance of emotions and FaceReader™

A Multiple Factor Analysis (MFA) (Escofier & Pages, 1994) was performed on the average dominance durations observed with TDE and average maximum facial intensities over tertiles observed with FaceReader™. Product spaces and correlation plots were constructed to visualize sample differences and similarities in emotion characteristics. RV coefficient was calculated from MFA analysis to investigate the correlation between FaceReader™ and TDE.

2.6.4. Liking scores

Mean liking scores and standard errors of the mean were calculated for the first, third and fifth bite per product. A three-way ANOVA was performed with product and bite as fixed factors and subject and its interaction effects with all fixed factors as random effects. A Tukey HSD pairwise comparison was performed upon significance of the ANOVA.

3. Results

3.1. Temporal dominance of emotions

Fig. 2 depicts the dominance bandplots for emotions per product for the first, third and fifth bite of consumption. All yogurt-granola samples were characterized by a dominance of *interested* feelings at the beginning of the first bite. The *hard:large:10%, hard:small:10%* and *hard:small:20%* were mainly characterized by *calm* and *good* feelings. Additionally, *hard:small:20%* was dominated by *enthusiastic* feelings at the beginning of the first bite and *happy* feelings at the beginning of the third bite. The *soft:large:10%, soft:small:10%* and *hard:small:3%* were mainly characterized by *calm* and *bored* feelings. The dominance duration of *interested* disappeared towards the fifth bite of consumption. Hardly any other dynamic changes could be identified for any of the other emotion descriptors between and within bites.

Table 3 shows the F-values of the ANOVA on dominance durations in % of standardized time for each attribute by product, bite and tertile obtained with TDE. The significant interaction effect of bite by tertile ($F(4,2420) = 4.8$, $p < 0.001$) indicates that the dominance durations of *interested* feelings significantly decreased from the beginning to the end of a bite, but that these dynamic changes were specific for the first and third bite. A main bite effect was observed for *interested* ($F(2,110) = 19.0$, $p < 0.001$), which shows that dominance durations of *interested* feelings significantly decreased from the first to the fifth bite for all products. Significant interaction effects for product by tertile were observed for *bored* ($F(10,2420) = 2.8$, $p = 0.002$), *energetic* ($F(10,2420) = 2.1$, $p = 0.02$) and *happy* ($F(10,2420) = 1.9$, $p = 0.04$), meaning that the dominance durations of these attributes did not develop the same way over tertiles between products. *Bored* feelings significantly increased in dominance duration from the first to the third tertile for the *hard:small:3%,* but no significant effect between products and tertiles were observed for *energetic* and *happy* when performing Tukey HSD pairwise comparison. Significant product by bite interaction effects were observed for *calm* ($F(10,2420) = 3.7$, $p < 0.001$), *disgusted* ($F(10,2420) = 3.5$, $p < 0.001$), *enthusiastic* ($F(10,2420) = 2.9$, $p = 0.001$), *good* ($F(10,2420) = 2.6$, $p = 0.004$) and *whole* ($F(10,2420) = 1.9$, $p = 0.04$), which indicates that the dynamic changes in dominance durations between bites were product specific. From first to fifth bite, *calm* feelings significantly increased in *soft:small:10%, disgusted* feelings significantly decreased in *soft:big:10%* and *enthusiastic* feelings significantly decreased in *hard:small:20%*. No significant effects between products and bites were observed for *good* and *whole* after pairwise comparison using Tukey HSD.

3.2. Dynamic facial expressions

Table 4 shows the ANOVA results of the maximum facial expression intensities by product, bite and tertile observed with FaceReader™.
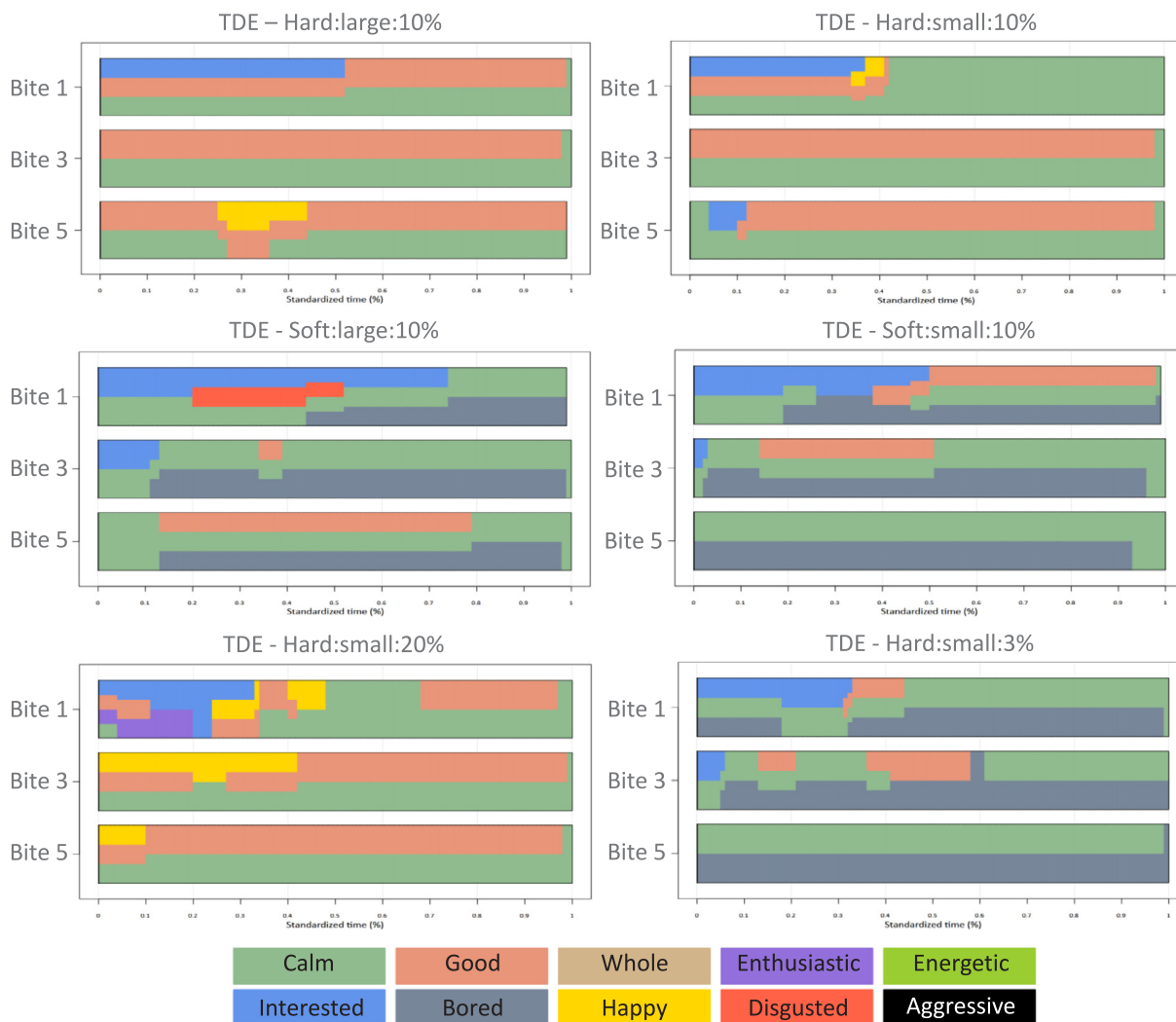
Fig. 2. TDE bandplots of the sequence of dominant emotions by product for the first, third and fifth bite for all six yogurt-granola products. Coloured rectangles represent the dominant emotions and are stacked at each moment, displaying multiple dominances (without taking into account dominance rates) at a given time point. Represented emotions are significant at p < 0.05.

Table 3
ANOVA of dominance durations (in % of standardized time) by product, bite, tertile and its interactions observed in Temporal Dominance of Emotions.

Descriptor	F_{prod}	F_{bite}	$F_{tertile}$	$F_{prod*bite}$	$F_{prod*tertile}$	$F_{bite*tertile}$
Aggressive	**2.5***	1.2	0.8	0.8	0.5	0.3
Bored	**15.2***	2.1	**5.7**	1.5	**2.8**	0.7
Calm	0.7	**5.0**	**3.3***	**3.7***	0.2	0.1
Disgusted	**5.5***	1.7	2.9	**3.5***	1.2	0.1
Energetic	**5.0***	0.7	0.1	1.1	**2.1***	0.3
Enthusiastic	**12.2***	**4.9**	2.5	**2.9**	1.3	1
Good	**5.8***	**3.6***	0.6	**2.6**	0.6	0.9
Happy	**7.3***	1.2	2.4	1.3	**1.9***	1
Interested	0.6	**19.0***	**12.6***	1.7	0.3	**4.8***
Whole	0.9	2.2	2.3	**1.9***	1.7	2.1

F-values in bold are significant at (*) 0.05, (**) 0.01, (***) 0.001.

Products could be differentiated based *neutral* (F(5,275) = 3.8, p = 0.002), *angry* (F(5,275) = 3.5, 0.004), *sad* (F(5,275) = 3.2, p = 0.009), *surprised* (F(5,275) = 2.6, p = 0.03) and *bored* (F (5,275) = 5.0, p < 0.001) facial expressions. Significant differences in facial expressions between products are visualized in Fig. 3. *Hard:-large:10%* was characterized by highest *neutral* facial expression intensities and lowest *sad* facial expressions. *Soft:large:10%* and *soft:-small:10%* were characterized by highest *angry* and *sad* facial expression

Table 4
ANOVA of maximum facial expression intensities by product, bite, tertile and its interactions.

Facial expression	$F_{product}$	F_{bite}	$F_{tertile}$	$F_{prod*bite}$	$F_{prod*tertile}$	$F_{bite*tertile}$
Neutral	**3.8**	**7.8***	**14.2***	1.8	**2.5**	0.5
Angry	**3.5**	**8.6***	**16.9***	1.7	0.8	0.5
Contempt	1.8	2.1	1.2	**3.5***	1.6	0.8
Disgusted	0.6	1.9	**24.6***	1.4	0.8	0.4
Happy	0.8	0.9	2.7	**1.9***	0.7	0.2
Sad	**3.2**	1.7	**26.8***	**2.0***	1.2	1.5
Scared	1	1.6	**44.5***	1.4	0.9	0.3
Surprised	**2.6***	0.1	**14.6***	1.5	**2.2***	0.6
Interest	0.6	0.7	1.8	1.2	0.6	2.3
Bored	**5.0***	**3.8***	**28.8***	**2.5**	1.3	1.4
Confused	1.8	2.6	0.1	**2.2***	0.9	0.3
Arousal	0.9	1.7	**184.8***	0.5	1.2	0.7
Valence	1.8	0.7	**6.4**	**2.1***	1	0.4

F-values in bold are significant at (*) 0.05, (**) 0.01, (***) 0.001.

intensities, and had significantly lowest *neutral* facial expressions. *Hard:small:20%* had significantly highest *neutral* and *bored* facial expression intensities and lowest *angry* facial expressions.

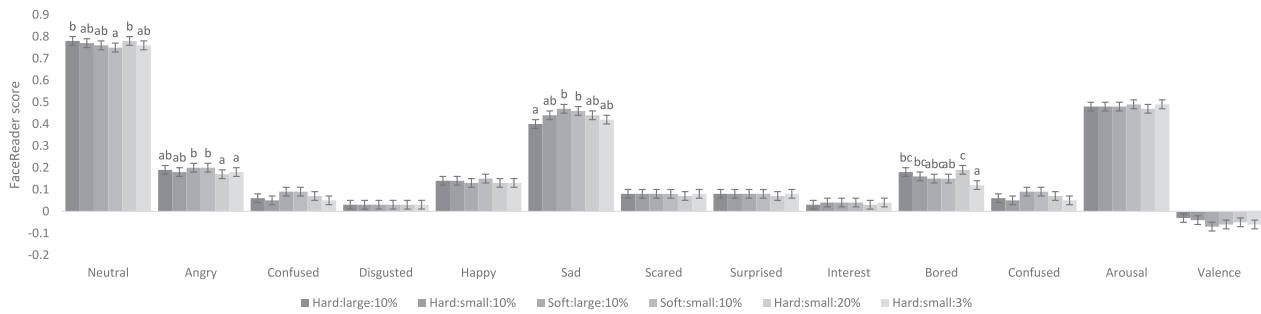Looking at dynamic changes between bites, we observed that *neutral*

Fig. 3. Graphic representation of the FaceReader™ scores for each facial expression per product. Means with different letters indicate significant differences between products (p < 0.05).
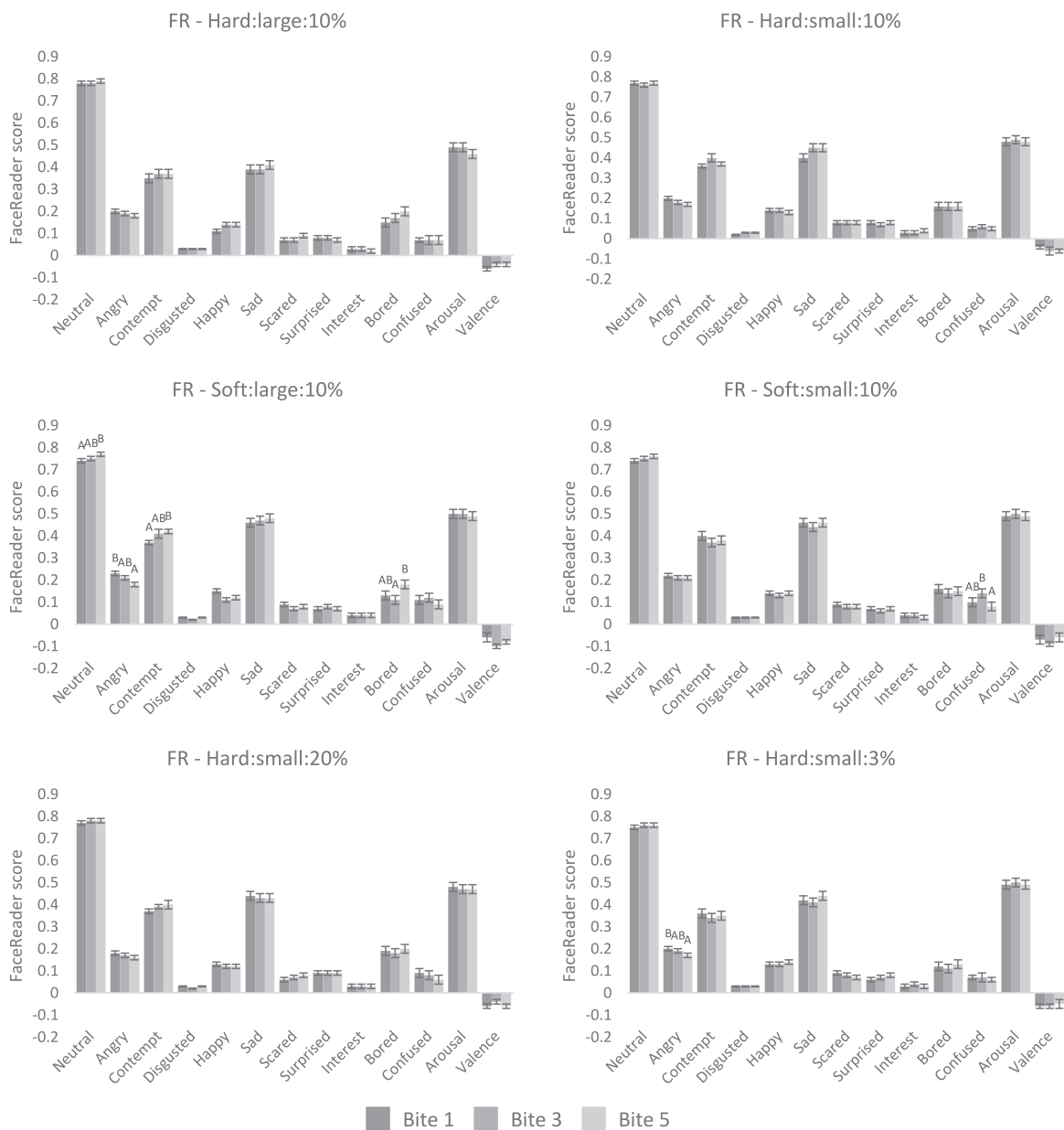

Fig. 4. Graphic representation of the FaceReader™ scores of the facial expressions for the first, third and fifth bite for all six yogurt-granola products. Means with different letters indicate significant differences between bites within a product (p < 0.05).

(F(2,110) = 7.8, p < 0.001) facial expressions significantly increased, and *angry* (F(2,110) = 8.6, p < 0.001) facial expressions significantly decreased from the first to the fifth bite of consumption for all products (Table 4). However, these significant main effects for the dynamic changes between bites for *neutral* and *angry* facial expressions were driven by changes in facial expressions for *neutral* and *angry* for *soft:large:10%* and *hard:small:3%*. Product by bite interaction effects indicated that product specific changes in facial expressions which were observed for *contempt* (F(10,2415) = 3.5, p < 0.001), *happy* (F (10,2415) = 1.9, p = 0.04), *sad* (F(10,2415) = 2.0, p = 0.03), *bored* (F (10,2415) = 2.5, p = 0.005), *confused* (F(10,2415) = 2.2, p = 0.01) and *valence* (F(10,2415) = 2.1, p = 0.03). Fig. 4 shows the significant changes in facial expressions per product for the first, third and fifth bite of consumption. No significant change between bites for any of the facial expressions observed with FaceReader™ were seen for *hard:-large:10%, hard:small:10%* and *hard:small:20%*. *Soft:small:10%* revealed most dynamic changes in facial expressions over bites, such as the significant increase of *neutral, angry, contempt* and *bored* facial expressions and a significant decrease in *angry* facial expressions from the first to the fifth bite. Moreover, *angry* facial expressions decreased from the first to the fifth bite for *hard:small:3%* and *confused* facial expressions decreased from the third to the fifth bite for the *soft:small:10%*. Posthoc analysis did not reveal significant differences between within products for *happy* facial expressions and *valance*.

Main tertile effects indicate the dynamic change of facial expressions within bites. Significant main tertile effects were observed for *angry* (F(2,110) = 16.9, p < 0.001), *disgusted* (F(2,110) = 24.6, p < 0.001), *scared* (F(2,110) = 44.5, p < 0.001) and *arousal* (F (2,110) = 184.8, p < 0.001), indicating that these facial expressions significantly decreased from the beginning to the end of each bite for all products. Interaction effects for product by tertile showed that the dynamic changes from beginning to the end of a bite for *neutral* (F (10,2414) = 2.5, p = 0.006) and *surprised* (F(10,2525) = 2.2, p = 0.02) facial expressions were product specific. *Neutral* facial expressions decreased from beginning to end of each bite for *hard:-small:10%, soft:large:10%, soft:small:10%* and *hard:small:20%*, and *surprised* facial expressions decreased from beginning to end of each bite for *hard:large:10%, hard:small:10%* and *hard:small:20%*.

3.3. Multivariate comparison of Temporal Dominance of Emotions and facial expressions

Fig. 5 shows the MFA plot which indicates product differentiation for the first, third and fifth bite based on an attribute's dominance durations observed with TDE (green font) and maximum facial expression intensities observed with FaceReader™ (red font). The MFA correlation circle (Fig. 5a) visualizes the emotion attributes in TDE and FaceReader™. The MFA individual factor map (Fig. 5b) represents the six products in black as mean points and the emotion configurations of the emotion measures in colour. The first two dimensions account for 57% of the variance (42.2% and 14.9% respectively). Products are discriminated along the first dimension, which reflects both valence and arousal, and differentiates the products from least liked (*soft:-small:10%, soft:large:10%* and *hard:small:3%*), to moderately liked (*hard:small:10%* and *hard:large:10%*) to most liked (*hard:small:20%*). The horizontal reflection of TDE and FaceReader™ emotions limits product differentiation of the products along a single dimension. Consumers self-reported mainly high arousal (*energetic* and *enthusiastic*) and positive (*happy, whole* and *good*) emotions and expressed *surprised, bored* and *neutral* facial expressions for the *hard:large:10%, hard:small:10%* and *hard:small:20%*. Least liked products were mainly characterized by low arousal (*bored* and *calm*) and negative (*disgusted* and *aggressive*) emotions using TDE. FaceReader™ further discriminates the least liked products by separating the soft:large:10% from the soft:small:10% along the second dimension. The *soft:large:10%* was mainly characterized by *sad, confused* and *interested* facial expressions, whereas

soft:small:10% and *hard:small:3%* were characterized by negative (*angry, disgusted* and *scared*) and *happy* facial expressions.

A significant RV coefficient of 0.545 (p < 0.001) was observed, representing a moderate correlation between the product configurations defined by the implicit (FaceReader™) and explicit (TDE) emotion measures. Overlapping emotion terms in both methods such as *happy* and *bored* seem negatively correlated, indicating that they are likely to have different meanings in TDE and FaceReader™. *Bored* observed with FaceReader™ seems positively correlated to positive (*happy* and *good*) and high arousal (*energetic* and *enthusiastic*) emotion terms in TDE. There seems to be more robustness on the agreement on negative emotion terms between TDE (disgusted, bored and aggressive) and FaceReader™ (disgusted, angry, confused and sad).

3.4. Dynamic liking

Fig. 6 shows the mean liking scores of the first, third and fifth bite of each product after TDE evaluations. Products could be differentiated based on their liking, whereas the *hard:small:20%* was significantly most liked followed by *hard:large:10%* and *hard:small:10%*, and the *soft:large:10%, soft:small:10%* and *hard:small:3%* were significantly least liked. A significant product by bite interaction effect (F10,550) = 3.4, p < 0.001) was observed for the liking scores after TDE evaluations, suggesting that liking scores did not evolve the same way for the six yogurt with granola samples of the three bites. The liking scores after TDE evaluations of the *hard:large:10%* significantly increased (p < 0.05) from the first to the fifth bite with 0.4. No other significant increase or decrease over bites was observed for any of the other products.

4. Discussion

This study compared the temporal evolvement of food-evoked emotions using a five-bite evaluation approach employing FaceReader™ and TDE. We hypothesized that the emotions obtained from facial expressions reflect the self-reported food-evoked emotion responses. Although FaceReader™ and TDE provide different type of information, we expected correspondence between FaceReader™ and TDE in terms of product discrimination and characterization (i.e. valence and arousal) within and between bites. Our findings indicate that FaceReader™ and TDE differentiate products differently and both methods show little overlap regarding type of emotion characterization. FaceReader™ and TDE discriminated products mainly along the valence dimension (positive – negative), which directly reflected product discrimination in terms of liking. Furthermore, food-evoked emotion profiles obtained with FaceReader™ and TDE show little dynamics within and between bites.

Consumers mainly self-reported positive (*good*) and low arousal (*calm* and *bored*) feelings using TDE, while highest intensities for *neutral, arousal* and negative (*sad* and *contempt*) facial expressions were observed using FaceReader™. Similar emotion terms in TDE and FaceReader™, such as *happy, bored, interested* and *disgusted*, do not seem to have similar meanings in both methods. We observed that *happy* facial expressions are negatively correlated to subjective *happy* feelings reported with TDE. Moreover, *happy* facial expressions are correlated to negative emotion terms, such as *angry, scared* and *disgusted*. Danner, Haindl, Joechl, and Duerrschmid (2014) reported similar findings and suggests that the detection of *happy* facial expressions by FaceReader™ needs more expressive facial movements (e.g. smiling), which could be hampered by the individual assessments of foods in laboratory settings and lack of social interactions that invites people to be more articulating and expressive of their facial movements.

Least liked products (*hard:small:3%, soft:small:10%* and *soft:-large:10%*) were associated with negative emotions and most liked products (*hard:small:20%, hard:small:10%* and *hard:large:10%*) were characterized by positive emotions. FaceReader™ further differentiated
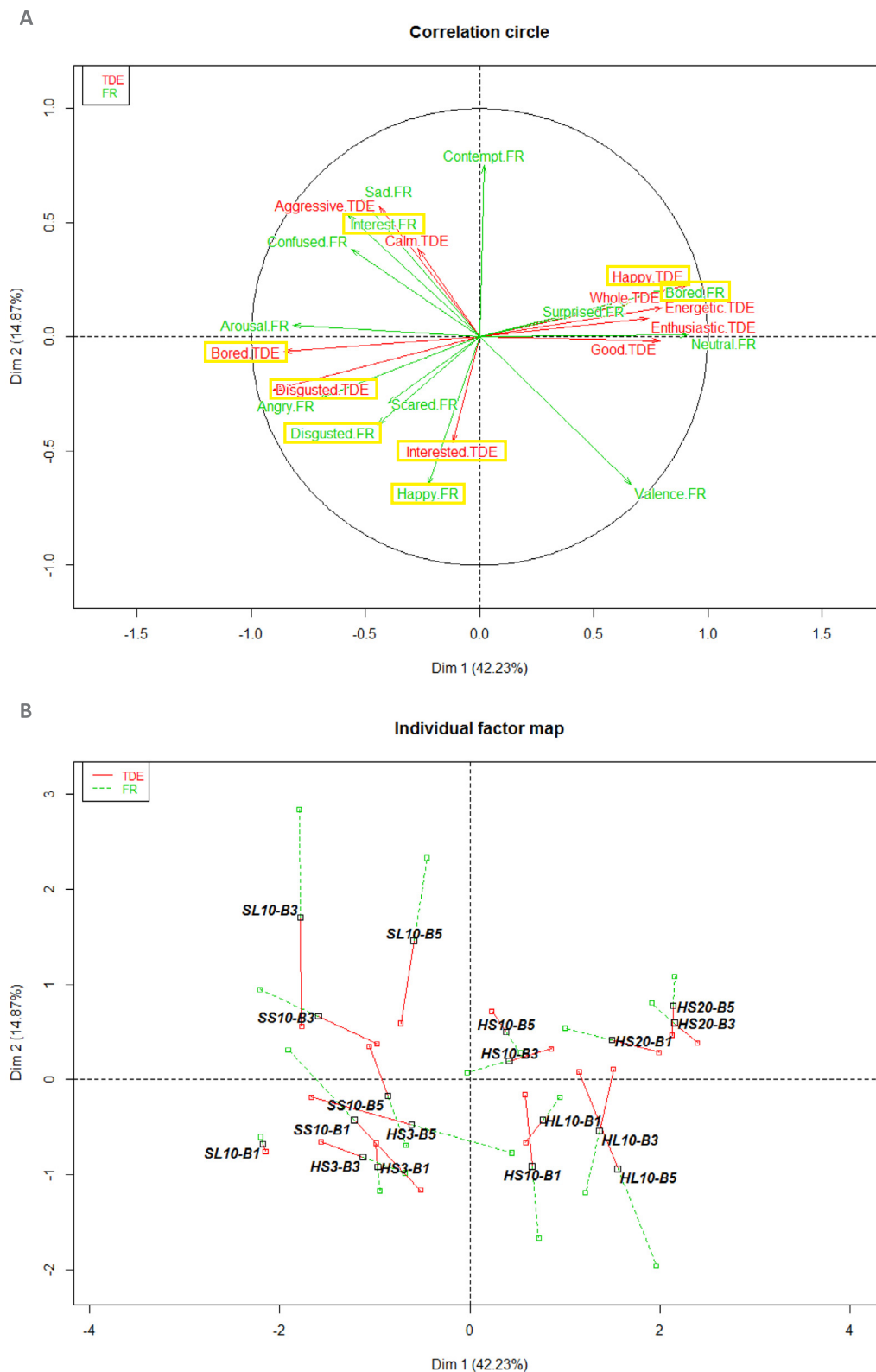
Fig. 5. Representation of the first two dimensions of the MFA space showing (a) emotion descriptors based on mean dominance durations observed with TDE (green font) and mean maximum facial expression intensities observed with FaceReader™ (red font), and (b) the first (B1), third (B3) and fifth (B5) bite of the six products as mean points, the partial individuals representing the emotion configurations of the products of the two emotion measurements. HS10 (hard:small:10%), HL10 (hard:large:10%), SS (soft:small:10%), SL10 (soft:large:10%), HS20 (hard:small:20%), HS3 (hard:small:3%). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
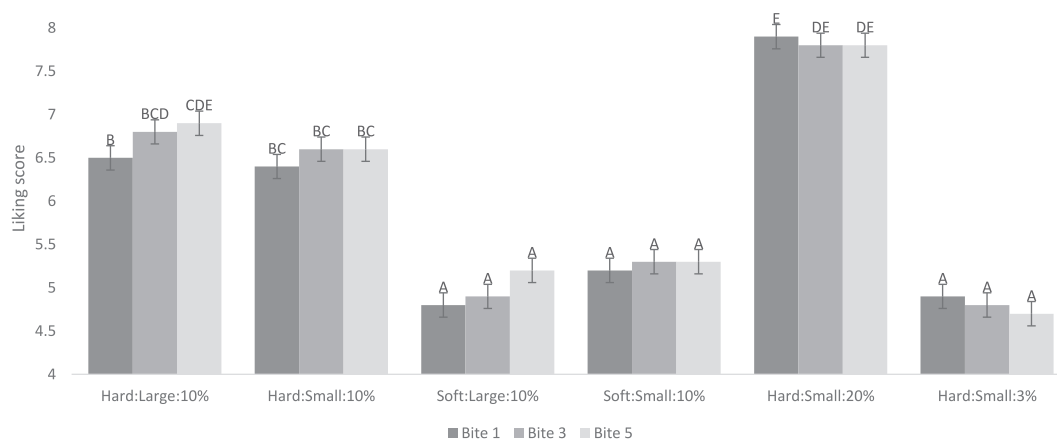
Fig. 6. Mean liking scores and standard errors of the mean by product and bite after TDE evaluations. Means with different letters indicate significant differences between products and bites (p < 0.05).

the least liked products from each other on arousal and negative facial expressions. These findings are in line with previous research that suggests that facial expressions are more suitable to characterize and differentiate disliked products compared to liked products (Zeinstra et al., 2016, Danner, Sidorkina, Joechl, & Duerrschmid, 2014).

In line with previous research, we observed that negative facial expressions were more intense than positive facial expressions (Danner et al., 2014; de Wijk, Kooijman, Verhoeven, Holthuysen, & de Graaf, 2012; Rocha-Parra, García-Burgos, Munsch, Chirife, & Zamora, 2016). Zeinstra, Koelen, Colindres, Kok, and De Graaf (2009) suggested that facial expressions are more suitable to measure dislikes than likes because negative facial expressions are quicker to appear and less influenced by other factors compared to positive facial expressions. FaceReader™ was originally developed for consumer products other than foods, hence the type of facial expression terms in FaceReader™ are skewed towards negative emotions. To steer product development and to tailor products to consumer's preferences, food-evoked emotion research targets regular product consumers. Regular product consumers mainly have positive emotion responses to products (so-called hedonic asymmetry), compared to non-users who have more negative or no emotion responses (King & Meiselman, 2010; Schifferstein & Desmet, 2010). This raises the question whether facial expression analysis will provide the desirable product information needed to steer product development.

Consumers self-reported *interested* feelings upon the first encounter of the product (e.g. beginning of the first bite). Previous intrinsic and extrinsic product experiences of the same or similar products cause sensory and hedonic expectations (Fernqvist & Ekelund, 2014; Piqueras-Fiszman & Spence, 2015). It is plausible that taste perceptions in the first bite define taste expectations for the following bites of the same product, causing self-reported *interested* feelings to wear off towards to third and fifth bite of consumption.

Our results indicated that facial expressions were more dynamic within bites than between bites. Arousal and negative (*sad*, *scared* and *angry*) facial expressions significantly decreased from the beginning to the end of each bite and neutral facial expressions increased from beginning to end of each bite. Although FaceReader™ corrects for partial occlusion of the lower part of the face (e.g. when subjects put a spoon to their mouth) by Deep Face Classification method, we cannot exclude the possibility that changes in oral processing behaviour affect the observed changes in facial expressions. Consumers might have displayed different muscle activities during consumption due to oral processing behaviour. Consumers might have used different chewing motions during the initial processing while granola is still hard and change chewing motions towards swallowing at the end of a bite. More chewing movements are likely to display higher muscle activity or tension which could be recognized by FaceReader™ as higher intensities

of negative facial expressions, whereas swallowing a bite might reflect more relaxed facial muscles which could be interpreted as neutral facial expressions by FaceReader™. Consequently, products that require intense mastication or products with 'big' changes in oral processing from beginning to end of mastication might hamper the (correct) identification of facial expressions.

The present study observed some dynamic changes in facial expressions over bites, but the direction of change was inconsistent between products. In contrast to our results, Rocha-Parra et al. (2016) observed a significant decrease of negative facial expressions accompanied by a significant increase in liking from the first to the third sip for two red wines. We speculate that the difference in observed dynamics of food-evoked emotion responses over multiple bites is caused by a difference in reward value between yogurt with added granola and red wine. Red wine is considered a highly emotional product which is likely to provide high reward value compared to yogurt with added granola which is a more basic food product and is likely to provide low reward value. Consequently, the emotion response to yogurt with added granola remains more stable during consumption. Moreover, consumers appreciated the red wines more upon increasing number of sip (Rocha-Parra et al., 2016), whereas liking scores of our yogurt with added granola samples did not change from bite to bite. The type of product and hedonic changes during consumption might have driven dynamic differences in emotion response over multiple bite assessments.

The current study used multiple bite assessments, which has the advantage that it mimics more natural eating behaviours as consumers eat food portions with multiple bites. FaceReader™ identified more negative emotions, while TDE identified more positive emotions. Hence, implicit and explicit measurements seem to be a complementary option when it comes to profiling food-evoked emotions (Leitch et al., 2015; Rocha, Lima, Moura, Costa, & Cunha, 2019). From a methodological point of view, FaceReader™ and TDE have the advantage that they allow to record changes in a consumer's food-evoked emotion response over time. TDE allows descriptive profiling of consumers' subjective experience of food products, but is limited by the number of emotion terms that can be assessed at the same time. TDE includes a minimum of 8 and a maximum of 12 emotion terms (Jager et al., 2014). The balance in emotion terms (positive, negative, high arousal and low arousal) is of utmost importance. The limited number of descriptors included in TDE could have led to dumping effects. Moreover, TDE is based on the concept of dominance, allowing the selection of only one emotion term at a time compared which could lead to relevant loss of information on the dynamic food-evoked emotion perception of a consumer.

Recording facial expressions has the advantage that it captures fast changing emotions and targets the subconscious part of the emotion

experience. The downside of recording facial expressions is that it is prone to data loss due to technical failures (i.e. shadows in the face due to bad lightening or loss of eye contact with the camera) and coverage of the face (i.e. wearing glasses or having facial hair such as a beard or moustache). Recording facial expressions leads to a large data set, and screening, filtering and analysing the data is time consuming. Moreover, it is still unknown how oral processing affects the identification of facial expressions by FaceReader™. FaceReader™ technology uses a Deep Face Classification method which allows analysis of facial expressions when the lower part of the face is hidden. It is unclear how this potentially biases or limits the facial expressions when for example the recognition of a specific expression requires opening of the mouth. Moreover, type of product seems to have important implications and could enhance bias due to oral processing behaviour. Yogurt with granola varying in hardness, size and concentration could have hampered the identification of facial expressions due to the potentially different oral processing behaviours (e.g. yogurt with hard granola vs. yogurt with soft granola). To better understand how oral processing behaviour influences the identification of facial expressions, future research should be done with products from the same product category that evoke different oral processing behaviours (e.g. peach cubes vs. peach smoothie).

To conclude, the emotion profiles obtained with implicit measures (facial expressions) show little overlap with the emotion profiles obtained with explicit measures (TDE) due to different type and nature in descriptors. Food-evoked emotions were mainly mild and positive and emotion responses did not seem product specific, but rather relate to the product category. Food-evoked emotion responses were fairly stable from first to last bite and only very limited changes were observed using implicit and explicit emotion measures. Both methods discriminated products mainly on the valence dimension (positive – negative) which directly reflected product discrimination in terms of liking.

Acknowledgements

References

Castura, J. C., Antúnez, L., Giménez, A., & Ares, G. (2016). Temporal Check-All-That-Apply (TCATA): A novel dynamic method for characterizing products. *Food Quality and Preference, 47*, 79–90.

Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. *IEEE Transactions on Pattern Analysis & Machine Intelligence,* (6), 681–685.

Danner, L., Haindl, S., Joechl, M., & Duerrschmid, K. (2014). Facial expressions and autonomous nervous system responses elicited by tasting different juices. *Food Research International, 64*, 81–90.

Danner, L., Sidorkina, L., Joechl, M., & Duerrschmid, K. (2014). Make a face! Implicit and explicit measurement of facial expressions elicited by orange juices using face reading

technology. *Food Quality and Preference, 32*, 167–172.

de Wijk, R. A., Kooijman, V., Verhoeven, R. H., Holthuysen, N. T., & de Graaf, C. (2012). Autonomic nervous system responses on and facial expressions to the sight, smell, and taste of liked and disliked foods. *Food Quality and Preference, 26*(2), 196–203.

Delarue, J., & Blumenthal, D. (2015). Temporal aspects of consumer preferences. *Current Opinion in Food Science, 3*, 41–46.

Escofier, B., & Pages, J. (1994). Multiple factor analysis (AFMULT package). *Computational Statistics & Data Analysis, 18*(1), 121–140.

Fernqvist, F., & Ekelund, L. (2014). Credence and the effect on consumer liking of food–A review. *Food Quality and Preference, 32*, 340–353.

Galmarini, M., Visalli, M., & Schlich, P. (2017). Advances in representation and analysis of mono and multi-intake Temporal Dominance of Sensations data. *Food Quality and Preference, 56*, 247–255.

Gutjar, S., Dalenberg, J. R., de Graaf, C., de Wijk, R. A., Palascha, A., Renken, R. J., & Jager, G. (2015). What reported food-evoked emotions may add: A model to predict consumer food choice. *Food Quality and Preference, 45*, 140–148.

He, W., Boesveldt, S., de Graaf, C., & de Wijk, R. A. (2016). The relation between continuous and discrete emotional responses to food odors with facial expressions and non-verbal reports. *Food Quality and Preference, 48*, 130–137.

Jager, G. (2016). Short-term time structure of food-related emotions: Measuring dynamics of responses. *Emotion Measurement* (pp. 273–298). Elsevier.

Jager, G., Schlich, P., Tijssen, I., Yao, J., Visalli, M., De Graaf, C., & Stieger, M. (2014). Temporal dominance of emotions: Measuring dynamics of food-related emotions during consumption. *Food Quality and Preference, 37*, 87–99.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*(9), 697.

King, S. C., & Meiselman, H. L. (2010). Development of a method to measure consumer emotions associated with foods. *Food Quality and Preference, 21*(2), 168–177.

Köster, E. P. (2003). The psychology of food choice: Some often encountered fallacies. *Food Quality and Preference, 14*(5–6), 359–373.

Köster, E. P., & Mojet, J. (2015). From mood to food and from food to mood: A psychological perspective on the measurement of food-related emotions in consumer research. *Food Research International, 76*, 180–191.

Leitch, K., Duncan, S., O'keefe, S., Rudd, R., & Gallagher, D. (2015). Characterizing consumer emotional response to sweeteners using an emotion terminology questionnaire and facial expression analysis. *Food Research International, 76*, 283–292.

Mahieu, B., Visalli, M., Schlich, P., & Thomas, A. (2019). Eating chocolate, smelling perfume or watching video advertisement: Does it make any difference on emotional states measured at home using facial expressions? *Food Quality and Preference, 77*, 102–108.

Pineau, N., Schlich, P., Cordelle, S., Mathonnière, C., Issanchou, S., Imbert, A., ... Köster, E. (2009). Temporal Dominance of Sensations: Construction of the TDS curves and comparison with time–intensity. *Food Quality and Preference, 20*(6), 450–455.

Piqueras-Fiszman, B., & Spence, C. (2015). Sensory expectations based on product-extrinsic food cues: An interdisciplinary review of the empirical evidence and theoretical accounts. *Food Quality and Preference, 40*, 165–179.

Rocha-Parra, D., García-Burgos, D., Munsch, S., Chirife, J., & Zamora, M. C. (2016). Application of hedonic dynamics using multiple-sip temporal-liking and facial expression for evaluation of a new beverage. *Food Quality and Preference, 52*, 153–159.

Rocha, C., Lima, R., Moura, A., Costa, T., & Cunha, L. (2019). Implicit evaluation of the emotional response to premium organic herbal infusions through a temporal dominance approach: Development of the temporal dominance of facial emotions (TDFE). *Food Quality and Preference, 76*, 71–80.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161. https://doi.org/10.1037/h0077714.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information, 44*(4), 695–729.

Scherer, K. R. (2009). Emotions are emergent processes: They require a dynamic computational architecture. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1535), 3459–3474.

Schifferstein, H. N., & Desmet, P. M. (2010). Hedonic asymmetry in emotional responses to consumer products. *Food Quality and Preference, 21*(8), 1100–1104.

Schouteten, J. J., De Steur, H., Sas, B., De Bourdeaudhuij, I., & Gellynck, X. (2017). The effect of the research setting on the emotional and sensory profiling under blind, expected, and informed conditions: A study on premium and private label yogurt products. *Journal of Dairy Science, 100*(1), 169–186.

van Bommel, R., Stieger, M., Boelee, N., Schlich, P., & Jager, G. (2019). From first to last bite: Temporal dynamics of sensory and hedonic perceptions using a multiple-intake approach. *Food Quality and Preference*103748.

Zeinstra, G. G., Koelen, M., Colindres, D., Kok, F., & De Graaf, C. (2009). Facial expressions in school-aged children are a good indicator of 'dislikes', but not of 'likes'. *Food Quality and Preference, 20*(8), 620–624.