Food Quality and Preference 62 (2017) 60–70

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Sensory expectation, perception, and autonomic nervous system responses to package colours and product popularity

Aurelia Schulte-Holierhoek[a,*], Luz Verastegui-Tena[b], Robert P.G. Goedegebure[b], Betina Piqueras Fiszman[b], Paul A.M. Smeets[a,c]

[a] Division of Human Nutrition, Wageningen University & Research, Wageningen, Netherlands
[b] Marketing and Consumer Behaviour Group, Wageningen University & Research, Wageningen, Netherlands
[c] Image Sciences Institute, University Medical Center Utrecht, Utrecht, Netherlands

ARTICLE INFO

Keywords:
Autonomic nervous system
Heart rate
Skin conductance
Descriptive social norm
Packaging
Taste

ABSTRACT

Consumers' perception of, and behaviour towards, products are influenced by extrinsic cues, including packaging and social norms. However, the understanding of this process is unsatisfactorily captured by questionnaires. Autonomic nervous system (ANS) responses can be used to measure implicit consumer responses. The aim of this work was to assess how packaging cues and social norms influence product expectation, product perception, and ANS responses. Ninety-eight adults (age: 23.3 ± 3.2 years; BMI: 21.3 ± 2.2 kg/m^2) first viewed four images of a yogurt package modified in hue (blue/red), brightness (high/low), and saturation (high/low) and two dummies alongside a fictitious product popularity score. After each image presentation, participants rated their expectations of the yogurt, tasted, and rated their perception of it. Expectations and the perception of liking, healthiness, sweetness, and flavour intensity were rated on 100-unit VAS scales. Heart rate (HR) and skin conductance response (SCR) to the image and tasting were measured. The darker, saturated red package elicited the lowest expectation of healthiness and the highest expectation of flavour intensity and sweetness. Red packages increased SCR while blue packages decreased them. During yogurt tasting, low product popularity was associated with a stronger decrease in SCR than a high popularity. Overall, the measured ANS responses were small. In conclusion, this study was the first to look at the effect of expectations elicited by a product's packaging colour and popularity on explicit ratings and ANS responses. We found differences in SCR to package colour and product popularity, suggesting their importance in affecting consumer responses.

1. Introduction

New food products undergo extensive sensory and consumer testing before being launched (van Trijp & Schifferstein, 1995). One challenge when conducting these tests are the many factors that influence the final perception of the food and its flavour (Okamoto & Dan, 2013). These include human factors, such as the genetic background and current physiological state as well as food-related factors. Food-related factors play an important role in setting the frame of the hedonic and sensory product properties. These factors include the context, for example the location of sale, serving or consumption, as well as extrinsic product cues, such as brand, labelling, packaging colours and shapes (Okamoto & Dan, 2013; Piqueras-Fiszman & Spence, 2015). Therefore, it is not sufficient to evaluate a product solely on its hedonic or sensory properties without considering the interrelationship of these factors (Okamoto & Dan, 2013).

Extrinsic product cues are made up of information that is not physically part of the food but is related to the product. Their importance lies in the expectations that are elicited by them. Such expectations constitute *"pre-trial beliefs about the product"* and are based on knowledge and previous memories of the same or a similar cue (Okamoto & Dan, 2013). These expectations may influence consumer liking as well as sensory perception of the product itself (Piqueras-Fiszman & Spence, 2015). In other words: consumers deduct internal characteristics of a food including its taste from its external characteristics. In order to explain this effect, several models have been formulated, one of which is the 'assimilation-contrast' model (Piqueras-Fiszman & Spence, 2015). This model is based on the cognitive dissonance theory and states that the brain tries to avoid dissonance between expectation and experience by aligning the experience with the prior expectation (assimilation). Therefore, it is often observed that for small divergences the evaluation of the product shifts in the direction of

* Corresponding author.
E-mail addresses: a.schulte-holierhoek@gmx.net (A. Schulte-Holierhoek), luz.verasteguitena@wur.nl (L. Verastegui-Tena), robert.goedegebure@wur.nl (R.P.G. Goedegebure), betina.piquerasfiszman@wur.nl (B. Piqueras Fiszman), p.smeets@umcutrecht.nl (P.A.M. Smeets).

http://dx.doi.org/10.1016/j.foodqual.2017.06.017
Received 30 April 2017; Received in revised form 19 June 2017; Accepted 28 June 2017
Available online 29 June 2017
0950-3293/

the expectations. However, for large divergences, the consumer may magnify the discordance (contrast), shifting ratings in the opposing direction instead (Davidenko et al., 2015).

An extrinsic product cue that alters consumers' acceptance and expectations is the appearance of a food product. Several studies have assessed its effect focusing on particular characteristics, for example package colour. Piqueras-Fiszman and Spence (2011), showed that switching the colour of potato chips packaging (symbolising a different flavour), led consumers acquainted with the brand to report the switched or a wrong flavour. Similar effects were observed for milk desserts with respect to package shape and colouring, further stressing the importance of appropriate packaging in the product's appearance and acceptance (Ares & Deliza, 2010). Huang and Lu (2016) studied the effect of colour packaging on perceived healthiness and found that utilitarian products in blue packages were perceived as healthier than utilitarian products in red packages of the same brightness and saturation. Few studies have examined the differences in consumer response to colour manipulation while controlling for more than just the hue. An exception would be the study by Tijssen, Zandstra, Graaf, and Jager (2017) where packages differentiating in hue, brightness and saturations were used and whose findings include that brighter, less saturated colours lead to higher health ratings.

Another relevant determinant of products' acceptance stems from the social norms. Social norms influence us on a daily basis and can be defined as "*implicit codes of conduct that provide a guide to appropriate action*" (Higgs, 2015, p. 38). As such, we derive acceptable behaviour from social norms. One can distinguish between two different kinds of social norms, with varying effects on behaviour and attitude. Firstly, *injunctive norms* refer to endorsed, and expected behaviour in specific situations. *Descriptive norms,* on the other hand, describe the behaviour prevalent in the majority of a group. Whereas injunctive norms influence attitudes, descriptive norms strongly affect actual behaviour and lead to more short-term, automatic behavioural responses (Melnyk, van Herpen, & Trijp, 2010). Thus, descriptive norms appear to have a strong effect on immediate liking and consumption.

Previous research using descriptive social norms shows that people prefer the food they are led to believe is more popular (Robinson, 2015; Robinson, Fleming, & Higgs, 2014). Stok, Ridder, Vet, and Wit (2014) found that descriptive social norms increased the consumption of fruit in adolescents. Furthermore, Higgs (2015) found that consumers tend to adjust their liking scores to those of other present consumers. Likewise, Nook and Zaki (2015) showed that consumers receiving feedback on how others had apparently rated a food, tended to shift their liking rating towards the group's rating and the strength of this change correlated with reward-related brain activation in a food viewing task. As we tend to look at what most people do in many situations (Surowiecki, 2005), descriptive norms appear to be an important external influence that drives consumers' liking and consumption.

In sensory and consumer science, different psychophysiological measures have been used to assess attention effects. Eye tracking is one method often used to assess attentional information about a food, for example while modifying the product package (Fulcher, Dean, & Trufil, 2016; Piqueras-Fiszman, Velasco, Salgado-Montejo, & Spence, 2013). However, as useful as eye tracking is in documenting eye movements and characterising gazing behaviour and the visual attraction of food, it cannot measure the emotional state of the participant or the emotional valence of the stimulus (Hurley, Hutcherson, Tonkin, Dailey, & Rice, 2015; Mojet et al., 2015). Another potential psychophysiological measure that has received less attention is electroencephalography. Studies have shown that it can be used to show differences in cortical activation between high and low calorie food exposure (Ohla, Toepel, Le Coutre, & Hudry, 2012), as well as liked and disliked brands (Bosshard et al., 2016). Further research into the measure has been successful (Martin, 1998), however its potential is limited by the fact that only cortical reactions can be measured, limiting it to approach and avoidance related emotions (Spinelli & Niedziela, 2016; Wijk & Boesveldt, 2016).

A further approach used to assess the influence of extrinsic cues on consumers involves measuring responses of the Autonomic Nervous System (ANS) (Wijk & Boesveldt, 2016; Wijk, He, Mensink, Verhoeven, & Graaf, 2014). The ANS controls autonomic bodily processes such as heart rate (HR), respiration and digestion. By measuring ANS responses, it is possible to collect information on reactions of the body to external stimuli, such as food images, thus making it possible to assess their emotional impact (Kreibig, 2010). Especially in cases where no significant difference in liking between products is found with traditional methods, implicit responses, such as those recorded by ANS responses, may help to assess future product performance (Spinelli & Niedziela, 2016). Furthermore, ANS responses allow for the simultaneous recording of different real time, non-invasive parameters (Rousmans, Robin, Dittmar, & Vernet-Maury, 2000).

The ANS measures recently employed in sensory science that have led to a higher discriminating potential among sensory stimuli and more robust results are HR and skin conductance. HR is controlled by a combination of sympathetic and parasympathetic activity (Wijk & Boesveldt, 2016) and is associated with arousal (Danner, Haindl, Joechl, & Duerrschmid, 2014). Skin conductance is controlled by the sympathetic nervous system and varies depending on the activity of the sweat glands in the skin with arousal of any kind leading to a higher activity of the glands. Skin conductance includes both the background tonic, or skin conductance level (SCL) and skin conductance response (SCR), which are rapid phasic components that result from sympathetic neuronal activity (Spinelli & Niedziela, 2016).

Only a few studies have conducted research on the ANS responses to food exposure with inconsistent results. Nederkoorn, Smulders, and Jansen (2000) found an increase in HR and SCR compared to baseline in females exposed to liked foods (Wijk, Kooijman, Verhoeven, Holthuysen, & Graaf, 2012) found differences in ANS responses between liked and disliked foods in children and adults. However, Hurley et al. (2015) could not confirm these findings in a study using cereal packages, indicating that the differences found by Nederkoorn et al. (2000) and Wijk et al. (2012) might be due to odour and not visual exposure. Therefore, further research is needed to properly assess the effect of packaging cues on ANS responses.

The majority of studies on ANS responses have focussed on taste. Rousmans et al. (2000) found that ANS responses to sweet taste (pleasant) were the weakest, whereas the response to bitter taste (unpleasant) stimuli was the strongest. Research with juices found significant differences in SCR, but not HR between different juices ranging from sauerkraut (salted cabbage fermented in brine with a sour/salty taste) to mixed vegetable, grapefruit, orange, and banana juice (Danner et al., 2014). In addition, HR and skin temperature responses were shown to differentiate between five similarly liked breakfast drinks (Wijk et al., 2014). While these studies have evaluated the validity of ANS responses in differentiating between similar foods, there is no research on the extent to which ANS responses are affected by cognitive and sensory expectations formed by extrinsic product properties such as packaging and social norms.

The primary objective of this study is to assess how a change in expectations caused by manipulating extrinsic food properties alters consumers' explicit (ratings) and implicit responses (ANS) to a food. The secondary objective is to study the effect of descriptive social norms on explicit and implicit responses to food. Our specific research questions were: (1) How do manipulations of hue, brightness, and saturation of a yogurt package change the explicit expectation and perception of the attributes healthiness, flavour intensity, and sweetness, as well as implicit ANS responses HR and SCR to the product? And (2) how does the descriptive social norm product popularity change liking, HR, and SCR?

Regarding the first research question, it was hypothesised that blue packages would be rated as healthier than red ones (Huang & Lu, 2016) and that brighter versions would be rated as healthier than darker versions (Mai, Symmank, & Seeberg-Elverfeldt, 2016). It was hypothesised that expected and perceived sweetness and flavour intensity would be higher for red packages. Additionally, it was expected that packages with lower brightness and higher saturation would increase expected flavour intensity as well as expected sweetness (Tijssen et al., 2017). Furthermore, it was expected that while viewing the yogurt package image HR would increase more compared to baseline for the red packages and SCR would decrease more for packages with a high saturation, independent of hue (Jacobs & Hustmyer, 1974; Wijk & Boesveldt, 2016; Zielinski, 2016). In addition, it was expected that while tasting the yogurt, HR would increase and significantly vary between colours. Additionally, we hypothesised that SCR would be positive during tasting and would not vary significantly between colours (Wijk et al., 2014).

With regard to the second question, it was expected that a high popularity score would lead to an increased hedonic rating of the yogurts, compared to those paired with a low popularity score (Higgs, 2015; Nook & Zaki, 2015). As consumers tend to choose in line with the group, we expected that a high popularity increases product expectations, reflected by a stronger ANS response to the product image and informed tasting indicating arousal.

2. Materials and methods

2.1. Participants

One hundred students and staff members of the University of Wageningen were recruited for this study. Ninety-eight eligible subjects completed all parts (mean age: 23.3 ± 3.2 years with a mean BMI of 21.3 ± 2.2 kg/m^2). Subjects were eligible if they consumed yogurt at least six times a year and were not colour blind, currently smoking, intolerant or allergic to common yogurt ingredients or had a known heart condition. Participants were recruited via flyers, posters, and social media. All participants provided written informed consent and received monetary compensation for their participation.

Participants were randomly assigned to either one of two product popularity conditions (Group A or B). Each group viewed each of the four colour modified packages, while the displayed product popularity score was counterbalanced across the two groups. For example, in group A participants viewed the red, high brightness, low saturation package with a high popularity (74%), whereas group B was shown a low popularity (26%) and the other way around for the red, low brightness, high saturation package, as shown in Table 1. The participant characteristics, as shown in Table 2, did not differ between groups, with the exception of the Attention to Social Comparison Information (ATSCI) score.

Table 1
Package colour and popularity scores displayed during the experiment by group.

Package:	Group:	
	A[1]	B[1]
BHL	Non-popular 24%	Popular 76%
BLH	Popular 76%	Non-popular 24%
RHL	Popular 74%	Non-popular 26%
RLH	Non-popular 26%	Popular 74%
Dummy 1	Non-popular 28%	Popular 73%
Dummy 2	Popular 73%	Non-popular 28%

[1] Participants were randomised into either one of the groups.

Table 2
Subject characteristics divided by group (mean ± SD).

	Total	Group A[b]	Group B[b]	p-value[c]
n	98	49	49	–
Gender: female (male)	58 (40)	29 (20)	29 (20)	–
Age (years)	23.3 ± 3.2	23.0 ± 2.9	23.5 ± 3.5	0.47[1]
BMI (kg/m^2)	21.3 ± 2.2	21.5 ± 2.5	21.1 ± 1.8	0.32[2]
Nationality: Dutch + Belgian (other)	52 (46)	27 (22)	25 (24)	0.69[3]
Dominant hand: right (left)	89 (9)	45 (4)	44 (5)	0.73[3]
Eating restraint (DEBQ[a])	2.4 ± 0.8	2.4 ± 0.8	2.3 ± 0.9	0.51[2]
General health interest score (HTAS[a])	4.5 ± 1.0	4.7 ± 0.9	4.4 ± 1.2	0.16[2]
ATSCI[a]	31.1 ± 8.0	32.8 ± 6.8	29.4 ± 8.9	0.03[2]
Interest in light products	3.7 ± 1.2	3.9 ± 1.3	3.5 ± 1.2	0.11[1]

[a] ATSCI: Attention to Social Comparison Information; DEBQ: Dutch Eating Behaviour Questionnaire; HTAS: Health and Taste Attitude Scale.
[b] Groups saw the same package stimuli, but with a different fictitious popularity score.
[c] P-value calculated with 1 Student's *t*-test, 2 Welch's *t*-test or 3 Chi-square test.

2.2. Experiment

2.2.1. Task structure

Participants were invited to take part in a study on the ANS responses to yogurt. Before the tasks, participants filled in a questionnaire. Participants then answered questions about their current thirst, hunger, satiety, desire to eat, prospective consumption, and fullness (according to Blundell et al., 2010). Subsequently, they completed the viewing and tasting task. An overview of the complete experiment is shown in Fig. 1.

To familiarise the participants with the procedure, a practice block preceded the experimental ones. The practice block included an example screen of a yogurt package image and its popularity, followed by three rounds of blind tasting (once with water and twice with the stimulus yogurt, without being shown a yogurt package before tasting). Afterwards the participants answered practice questions.

Participants completed six experimental blocks (Fig. 2), one for each modified package and the two dummies, each including a viewing and a tasting task (henceforth referred to as "informed tasting"). The blocks were completed with OpenSesame, version 3.1.2 (Mathôt, Schreij, & Theeuwes, 2012). Participants were instructed to keep as still as possible during the study, especially while viewing the yogurt packages and tasting the yogurt, to avoid possible artefacts.

Each block started with 60 s of relaxation. Afterwards participants viewed the package image together with the fictitious product popularity and subsequently answered questions about their expectations of the product. Once the participants had answered the questions about the package image, they tasted the yogurt. For each tasting participants were given approximately 27 g of yogurt. The sample was held at chin level in front of the participant. The screen instructed the participants when to take the straw into their mouth, when to take one sip of the yogurt and when to swallow the yogurt. The screen instructing participants to take the straw into their mouth also included the presented yogurt package along with its popularity score. The amount consumed per tasting was recorded. After swallowing the yogurt, participants rated their product perception.

2.2.2. Extrinsic cue manipulation

The image stimulus used for this study was the package of a fat-free red fruit drinking yogurt without any added sugars (*Optimel 'puur rode vruchten' yogurt*), which was altered in hue (blue/red), brightness (high/low), and saturation (low/high), abbreviated: BHL, BLH, RHL, and RLH respectively (cf. Tijssen et al., 2017). Detailed characteristics of the packages in combination with the exact manipulation are shown in Table 3. Two dummy packages were included in the trials and modified in colour in order to hide the repetitiveness of the stimuli (package and identical yogurt) and thus mask the aim of the study.

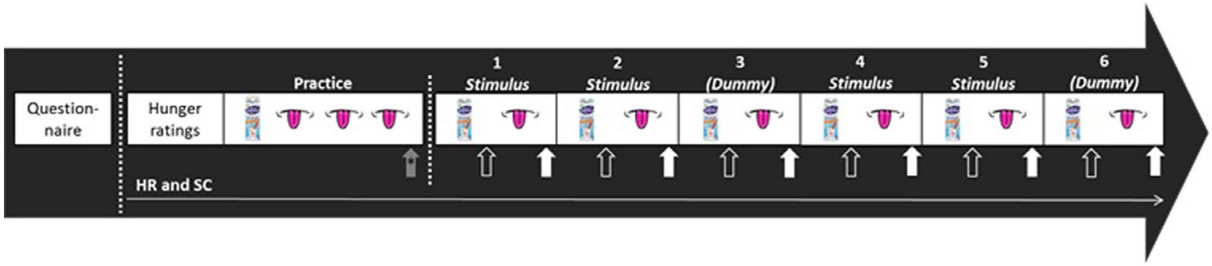

Fig. 1. Timeline of the experimental procedure. Participants completed two question blocks (questionnaire including DEBQ, HTAS, and ATSCI as well as their current hunger ratings), one practice and six experimental blocks. An experimental block was made up of seeing a colour modified yogurt package image (yogurt package), rating the package (black arrow), tasting the yogurt (tongue), and rating the taste (white arrow). Heart rate (HR) and skin conductance (SC) were recorded continually after the first questionnaire. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

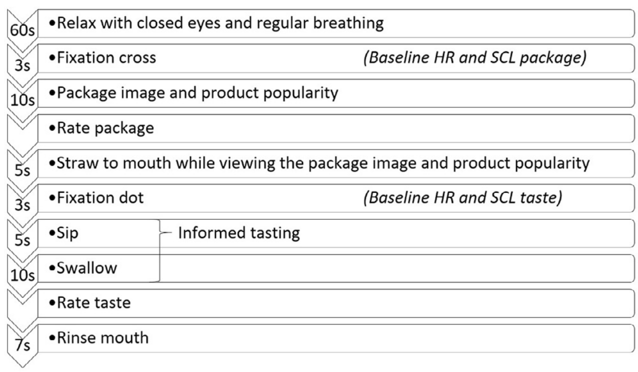

Fig. 2. Timeline of one experimental block. Participants completed six blocks, once for each of the four colour modified stimuli and the two dummies.

Vifit red fruit drinking yogurt ('*drinkyoghurt rode vruchten*') was used as the yogurt to be tasted for all four package stimuli. The dummy yogurt samples provided were Vifit raspberry drinking yogurt ('*drinkyoghurt framboos*') and Vifit strawberry drinking yogurt ('*drinkyoghurt aarbei*'). All three yogurt samples were plain white and visually indistinguishable with a similar viscosity.

2.2.3. Descriptive social norm manipulation

Each yogurt package was accompanied by a shopping basket symbol containing a fictitious percentage of people that, participants were told, had wanted to buy the yogurt in a previous study. The exact wording of the explanation to the participants was as follows: 'Additionally, you

Fig. 3. Screenshot illustrating a package example and the reported popularity of the yogurt.

Table 3
Characteristics of the colour modified packaging and yogurt provided with the corresponding package.

	BHL	BLH	RHL	RLH	Dummy 1	Dummy 2
						
Hue	NA	NA	NA	NA	−179	−121
Brightness[1]	+60 (high)	−15 (low)	+60 (high)	−15 (low)	−1	−10
Saturation[1]	−50 (low)	+30 (high)	−70 (low)	+30 (high)	+12	−3
Provided yogurt	Vifit red fruit drinking yogurt *('drinkyoghurt rode vruchten')*				Vifit strawberry drinking yogurt *('drinkyoghurt aarbei')*	Vifit raspberry drinking yogurt ('drinkyoghurt framboos'*)*

BHL: **B**lue, **h**igh brightness, **l**ow saturation; **BLH**: **B**lue, **l**ow brightness, **h**igh saturation; **RHL**: **R**ed, **h**igh brightness, **l**ow saturation; **RLH**: **R**ed, **l**ow brightness, **h**igh saturation.

[1] Brightness and saturation are provided relative to the original Optimel puur '*rode vruchten drin*k' package colour (blue hue).

will see a shopping basket next to the yogurt package. It shows how much consumers like the product. Liking scores were collected during a previous study by a PhD candidate.' Percentages of approximately 25% (i.e. low in popularity) and 75% (i.e. high in popularity) were chosen based on a pre-test showing that these numbers elicited the largest difference in responses. To support the displayed product popularity in a less cognitive way, the shopping baskets were coloured green to approximately 25% and 75%. An example screen is shown in Fig. 3.

2.2.4. Measures

The initial questionnaire included questions on participants' demographic information, pattern of yogurt consumption, the Dutch Eating Behaviour Questionnaire (DEBQ; 33 items, answered on a 5-point scale, including the sub-scale eating restraint; van Strien, Frijters, Bergers, & Defares, 1986), Health and Taste Attitudes Scale (HTAS; 38 items, answered on a 7-point scale, including the sub-scale general health interest; Roininen et al., 2001) and ATSCI (13 items, answered on a 6-point scale; Bearden & Rose, 1990). Participants rated the product expectations and perception on-screen on 100-unit visual analogue scales (VAS). They were asked about their liking of the product, how popular they thought the yogurt was after seeing the image (as a manipulation check), and about expected sweetness, flavour intensity, and healthiness. After tasting the yogurt, participants rated their liking and how popular they thought the yogurt was together with their perceived sweetness, flavour intensity, and healthiness again. The consumed amount of yogurt per block was measured with a digital scale and recorded in mg for each cup.

The VU University Ambulatory Monitoring System (VU-AMS, version 3.9; Willemsen, de Geus, Klaver, van Doornen, & Carroll, 1996) was used to measure HR and SCL. HR was measured in beats per minute (bpm) with seven electrodes (Ag/AgCl, Kendall ARBO H98SG ECG) placed on the chest. ECG was sampled at 1000 Hz. HR was calculated as the time between two adjacent R-waves. SCL was measured in μSiemens (μS) with two Ag-AgCl, non-polarizable electrodes filled with an isotonic electrode paste (0.05 M NaCl). The electrodes had a contact area of 6 mm and were attached to the index and middle finger of the non-dominant computer hand. SCL was sampled at 10 Hz with a signal range between 0 and 95 μS. The signal was filtered in a forward and reverse direction using a low pass filter with a cut-off frequency of 2 Hz.

2.3. Data analysis

HR and SCL were analysed with the use of the VU-AMS Data, Analysis & Management Software program (version 4.0). Suspicious heartbeats, indicating an artefact or abnormal heartbeat, were flagged by the program and manually inspected. All ectopic heartbeats were deleted. Additionally, measurements during which participants did not follow the timing or moved, as noted during the experiment, were excluded. The cleaned data was manually labelled between the automatically set markers from the VU-AMS plug-in for OpenSesame. The mean HR and SCL during fixation periods (3 s) before the package image and the yogurt tasting was subtracted from the following mean HR and SCR during the image and tasting, yielding the HR and SCR responses to the stimulus.

Statistical analysis was conducted in the R programming language and environment (R Core Team., 2013). A p-value ≤ 0.05 was considered significant. Mixed models were fitted using the *lmer* function of the *lme4* software package (Bates, Mächler, Bolker, & Walker, 2014) and analysed using the *anova* function in the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2016). Satterthwaits method for degree of freedom estimation was used. Tests for interactions were conducted. In the cases where an interaction was not significant, the interaction term was removed from the model. Sum of squares type II was used for models without interaction, whereas sum of squares type III was used for models with significant interaction terms. Post-hoc testing was conducted using the *lsmeans* package (Lenth, 2016), with the Tukey HSD method adjusting for multiple comparisons. To assess the fit of the models, underlying model assumptions were tested. Where the normality assumption seemed violated, permutation tests were conducted. In order to fulfil the assumption of normal residuals, the cube root of the baseline corrected package SCR data was used. After transformation, all assumptions were met for the analysed models.

To test for differences between the modified packaging colours, mixed model ANOVAs were conducted for expected and perceived healthiness, flavour intensity and sweetness each with *subject* as random variable and *package colour (BHL/BLH/RHL/RLH)*, *position (1/2/4/5)*, *gender (male/female)*, *hunger, light product interest*, and *nationality (Dutch and Belgian/other),* as fixed factors, as well as *sip size* for the perceived attributes. Interactions were tested between *package colour* and *position*, *package colour* and *displayed popularity score,* as well as *package colour* and *sip size* for perceived attributes. For the ANS measures HR and SCR, mixed model ANOVAs were conducted for the baseline corrected average while viewing the package and while tasting. *Subject* was considered a random factor, whereas *package colour*, *block position*, *gender, nationality,* and *hunger* were fixed factors. For the ANOVAs conducted for taste, *sip size* was also included as fixed factor. Interaction between *package colour* and *position*, as well as *package colour* and *sip size* was tested. To test for differences between the effect of a high and low popularity score, mixed model ANOVA of liking, HR and SCR were run with *subject* as random factor and *displayed popularity score (high/low)*, *ATSCI score,* and *group (A/B)* as fixed factors. For liking, the *package colour* along with the interaction between *popularity score* and *package colour* were also included. Spearman's correlation was used to assess the association between the explicit and implicit responses. *Gender*, *nationality*, *hunger*, *light product interest*, and *ATSCI score* were included in the data analysis to correct for their potential influence.

Due to technical difficulties with the ECG, as well as participants with previously unknown heart conditions, ANS measurements are missing for six participants. Furthermore, due to technical difficulties with the computer during four other experimental runs, only the hunger, hedonic, and sensory responses of the participants captured by the program are included in the analyses. Therefore, some responses are missing for these ten participants.

3. Results

3.1. Explicit responses

An overview of the explicit measurement results is given below. The *F* and *p*-values for the mixed models can be found in Table 4.

3.1.1. Healthiness

3.1.1.1. Package. There was an interaction effect for *package colour* and *block position* on the healthiness rating of the package image ($F(9, 328) = 4.17$, $p < 0.001$). As seen in Fig. 4(A1), healthiness varied, especially for RLH which showed a continuous decrease up to position 4, where RLH (39 ± 3) was rated significantly lower than the other three colours (BHL: 68 ± 4, BLH: 67 ± 3, RHL: 67 ± 3; all $p < 0.001$). There were significant main effects on the healthiness rating of the package image for both *package colour* ($F(3, 272) = 13.23$, $p < 0.001$) and *block position* ($F(3, 272) = 5.19$, $p = 0.002$). RLH (55 ± 2) was rated significantly lower than the other three packages (BHL: 66 ± 2, BLH: 66 ± 2, RHL: 64 ± 2; all $p \leq 0.001$), when averaging over *block position*. Furthermore, packages during the first block (66 ± 2) were rated significantly healthier than during the fourth (60 ± 2) and fifth block (59 ± 2; both $p \leq 0.050$).

3.1.1.2. Informed tasting. For the healthiness rating of the yogurt taste, an interaction effect between *package colour* and *block position* was found ($F(9, 291) = 2.06$, $p = 0.033$). Healthiness varied strongly between positions and colours, as seen in Fig. 4(A2). There was also a

Table 4
Results of the mixed model ANOVA for the explicit and implicit responses, and the descriptive social norm product popularity and liking.

		Healthiness						Flavour intensity						Sweetness					
Fixed factor		Package			Taste			Package			Taste			Package			Taste		
	df1	*df2*	*F*	*p*	*df2*	*F*	*p*	*df2*	*F*	*p*	*df2*	*F*	*p*	*df2*	*F*	*p*	*df2*	*F*	*p*
Colour	3	272	13.23	< 0.001	256	2.56	0.055	279	34.19	< 0.001	264	0.41	0.748	274	36.06	< 0.001	264	0.58	0.631
Block	3	272	5.19	0.002	256	0.69	0.558	279	0.52	0.670	264	4.16	0.007	274	0.52	0.669	264	0.39	0.762
Gender	1	90	0.68	0.413	86	2.33	0.131	91	0.68	0.413	87	2.50	0.118	92	0.23	0.635	87	0.10	0.759
Sip size	1	–	–	–	341	0.11	0.736	–	–	–	329	0.00	0.974	–	–	–	334	1.30	0.255
Hunger	1	93	1.92	0.169	88	5.33	0.023	93	1.17	0.282	88	0.18	0.673	94	0.20	0.659	88	0.19	0.668
Light product interest	1	93	6.73	0.011	88	7.47	0.008	92	0.24	0.626	87	0.25	0.621	94	0.00	0.999	87	0.01	0.923
Nationality	1	94	0.18	0.668	88	0.11	0.743	92	0.01	0.924	87	2.64	0.108	95	0.00	0.994	87	4.98	0.028
Colour * Block	9	328	4.17	< 0.001	291	2.06	0.033	–	–	–	–	–	–	334	1.98	0.041	–	–	–

		HR						SCR					
Fixed factor		Package			Taste			Package			Taste		
	df1	*df2*	*F*	*p*	*df2*	*F*	*p*	*df2*	*F*	*p*	*df2*	*F*	*p*
Colour	3	257	0.15	0.931	246	0.56	0.641	257	3.39	0.019	245	0.48	0.698
Block	3	257	0.99	0.398	246	1.86	0.137	257	2.16	0.093	246	1.76	0.155
Gender	1	87	3.10	0.082	83	0.26	0.611	89	0.07	0.794	84	0.21	0.645
Sip size	1	–	–	–	236	0.19	0.662	–	–	–	318	0.68	0.411
Hunger	1	89	0.96	0.330	85	2.02	0.159	90	0.29	0.588	85	0.49	0.485
Nationality	1	87	5.96	0.016	83	0.06	0.800	89	0.53	0.471	84	0.03	0.874

	Liking							HR						SCR					
Fixed factor		Package			Taste			Package			Taste			Package			Taste		
	df1	*df2*	*F*	*p*	*df2*	*F*	*p*	*df2*	*F*	*p*	*df2*	*F*	*p*	*df2*	*F*	*p*	*df2*	*F*	*p*
Popularity score	1	281	1.62	0.204	298	1.63	0.203	265	0.73	0.394	282	0.18	0.674	264	0.01	0.923	278	4.77	0.030
ATSCI score	1	93	4.86	0.030	94	2.57	0.112	93	0.04	0.839	90	0.08	0.783	94	0.00	0.962	88	0.29	0.593
Group	1	92	0.19	0.665	94	0.87	0.352	89	0.05	0.824	89	0.10	0.749	90	0.25	0.620	88	1.46	0.231
Colour	3	282	1.85	0.138	299	0.40	0.754	–	–	–	–	–	–	–	–	–	–	–	–
Popularity score * Colour	3	171	4.69	0.004	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Numerator (df1) and denominator (df2) degrees of freedom, as well as F and p-values are given for the fixed factors and interaction. – indicated that the factor was not included in the model

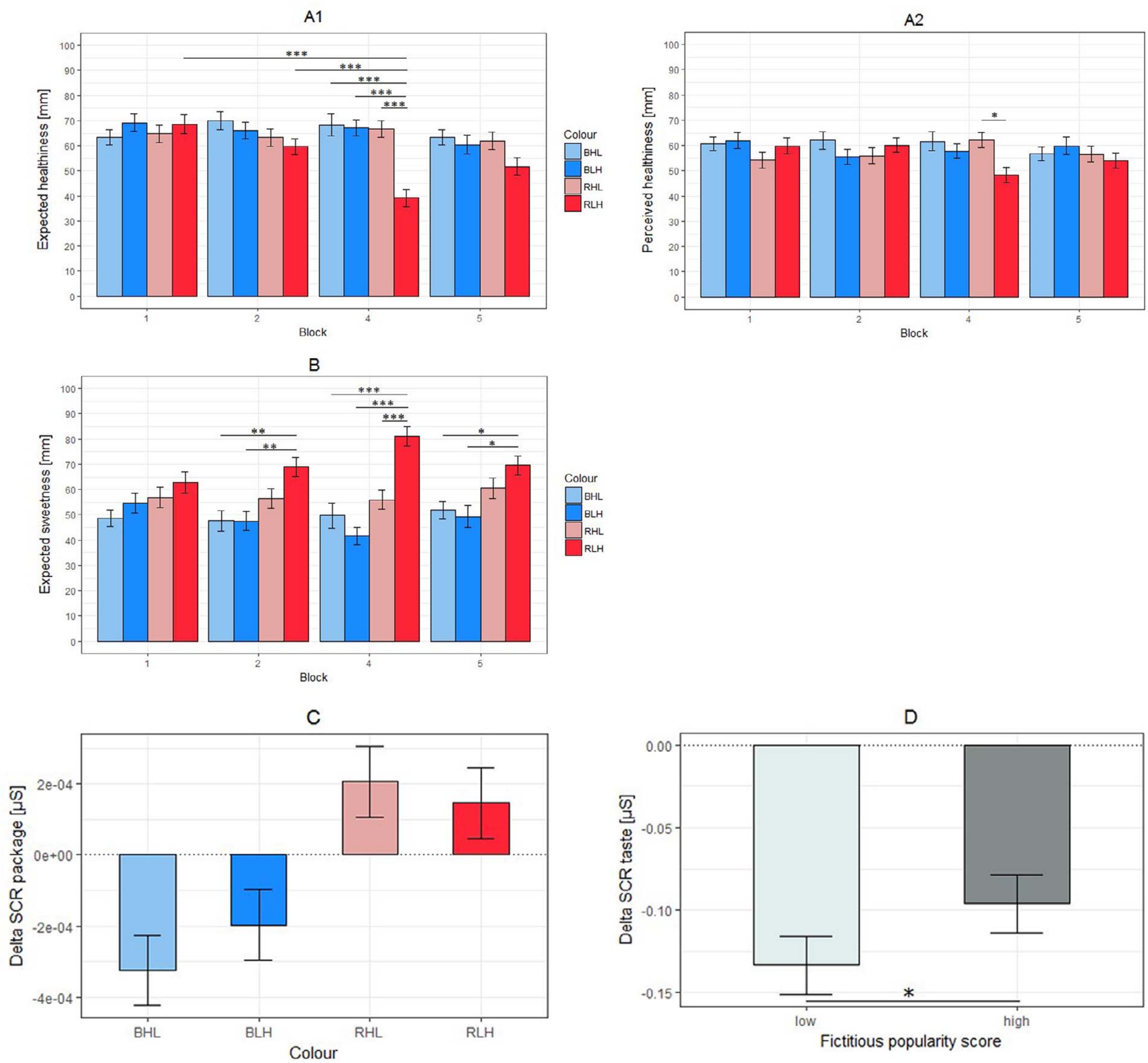

Fig. 4. Estimated marginal means by package colour or the interaction of package colour and block position depending on the results from mixed model ANOVA for the explicit and implicit attributes (A) expected (1) and perceived (2) healthiness, (B) sweetness, (C) skin conductance response (SCR), and estimated marginal mean by popularity score for (D) SCR the yogurt tasting. Displayed as estimated marginal means ± SE with significant differences within blocks and within colours signified by *(p < 0.05), **(p < 0.01) and ***(p < 0.001). An interaction effect was found for A1 and A2. BHL: Blue, high brightness, low saturation; BLH: Blue, low brightness, high saturation; RHL: Red, high brightness, low saturation; RLH: Red, low brightness, high saturation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

trend for a main effect of package colour ($F(3, 256) = 2.56, p = 0.055$), with RLH (56 ± 2) being rated as significantly less healthy than BHL (60 ± 2; $p = 0.044$), which follows the trend found for expected healthiness.

3.1.2. *Flavour intensity*

3.1.2.1. Package. For the expected flavour intensity, a main effect for package colour was found ($F(3, 279) = 34.19$, $p < 0.001$). RLH (72 ± 2) was rated significantly higher than BHL (52 ± 2; $t(279) = -8.86$, $p < 0.001$), BLH (53 ± 2; $t(279) = -8.56$, $p < 0.001$), and RHL (57 ± 2; $t(279) = -6.61$, $p < 0.001$).

3.1.2.2. Informed tasting. For the perceived flavour intensity, no main effect for *package colour* was found ($F(3, 264) = 0.41, p = 0.748$). Only a main effect on perceived flavour intensity was found for *block position* ($F(3, 264) = 4.16, p = 0.007$). Flavour intensity ratings increased with each position: A significantly lower rating was found for position 1 (56 ± 2) compared to position 5 (62 ± 2; $t(265) = -3.29$, $p = 0.006$).

3.1.3. *Sweetness*

3.1.3.1. Package. For expected sweetness, there was a significant interaction effect between *package colour* and *block position* ($F(9, 334) = 1.98$, $p = 0.041$). For the highly saturated colours RLH and BLH a difference was seen in positions 2, 4 and 5, where RLH was rated significantly higher than BLH on all positions (all $p \leq 0.050$). As seen in Fig. 4(B), the ratings for the other two colours, BHL and RHL, did not vary much between positions. *Package colour* also showed a main effect on expected sweetness ($F(3, 274) = 36.06$, $p < 0.001$). More precisely, RLH (71 ± 2) was rated as being significantly sweeter than the other package colours (BHL: 49 ± 2, BLH: 48 ± 2, RHL: 57 ± 2; all p < 0.001). Furthermore, RHL was rated as being significantly sweeter than the blue packages (both $p \leq 0.050$).

3.1.3.2. Informed tasting. No main effects on perceived sweetness were found. This includes p*ackage colour*, which did not influence the perceived sweetness of the yogurt ($F(3, 264) = 0.58$, $p = 0.631$). However, an effect of *nationality* was found with Dutch and Belgian participants perceiving the yogurt as sweeter than the other nationalities ($F(1, 87) = 4.87$, $p = 0.028$).

3.2. ANS responses

An overview of the implicit response results is given below. The *F* and *p*-values for the mixed model ANOVAs can be found in Table 4.

3.2.1. HR

3.2.1.1. Package. There was no main effect for *package colour* on delta HR while viewing the package image ($F(3, 257) = 0.15$, $p = 0.931$). However, a difference in HR was found depending on *nationality*, with Dutch and Belgian participants showing a stronger decrease in HR than the other nationalities ($F(1, 87) = 5.96$, $p = 0.016$).

3.2.1.2. Informed tasting. Package colour showed no main effect on delta HR ($F(3, 246) = 0.56$, $p = 0.641$).

3.2.2. SCR

3.2.2.1. Package. There was a main effect for *package colour* on delta SCR while viewing the package image, at $F(3, 257) = 3.39$, $p = 0.019$. As shown in Fig. 4(C), delta SCR increased for the red packages RHL (0.0002 ± 0.0001 μS) and RLH (0.0001 ± 0.000 μS) and decreased for the blue packages BHL (−0.0003 ± 0.0001 μS) and BLH (−0.0002 ± 0.0001 μS). However, post hoc testing did not yield significant differences in SCR between any of the package colours.

3.2.2.2. Informed tasting. No main effects on SCR while tasting the yogurt were found. Fixed factor *package colour* did not reach significance at $F(3, 245) = 0.48$, $p = 0.698$.

3.3. Descriptive social norm

An overview of the results is given below. The F and p-values of the mixed model ANOVAs can be found in Table 4. The popularity manipulation was considered successful, as the displayed popularity score had a significant effect on the rated popularity of the package ($t(379) = 10.90$, $p < 0.001$) and the taste ($t(380) = 8.49$, $p < 0.001$).

3.3.1. Liking

For the package image, liking scores depended on *package colour* and *popularity score* ($F(3, 171) = 4.69$, $p = 0.004$). Within the low popularity score, a significant difference was found between BHL (66 ± 6) and RLH (52 ± 6; $t(282) = 3.43$, $p = 0.016$). No significant effects of the taste perception on liking were found.

3.3.2. HR

There was no significant main effect for *popularity score* while viewing the package ($F(1265) = 0.73$, $p = 0.394$) or while tasting ($F(1, 282) = 0.18$, $p = 0.674$).

3.3.3. SCR

While viewing the package no significant main effect for *popularity score* was found ($F(1, 264) = 0.00$, $p = 0.923$). However, there was a significant main effect for *popularity score* while tasting the yogurt ($F(1, 278) = 4.77$, $p = 0.030$). A stronger decrease in SCR for the packages displayed with a low popularity score (−0.133 ± 0.018 μS) was found compared to those displayed with a high popularity score (−0.096 ± 0.018 μS; $t(278) = 2.19$, $p = 0.030$), as shown in Fig. 4(D).

3.4. Correlations implicit and explicit responses

Weak correlations were found between SCR and expected flavour intensity ($r = 0.13$, $p = 0.012$) as well as perceived healthiness ($r = -0.13$, $p = 0.014$), perceived flavour intensity ($r = -0.12$, $p = 0.018$), and liking of the taste ($r = -0.12$, $p = 0.020$).

4. Discussion

This study was the first to examine the effects of package colour and social norms on expectations, informed sensory taste perception ratings, and ANS responses. As a main finding, colour had an effect on SCR. The red packages elicited an increase in SCR, while the blue packages elicited a decrease in SCR. Additionally, in the informed tasting, product popularity exerted a significant effect on SCR, with a low popularity score leading to a stronger decrease in SCR compared to a high popularity score. None of the manipulations showed an effect on HR. Significant differences between package colours were found for the expectations of healthiness, flavour intensity, and sweetness. The red package with low brightness and high saturation (RLH) was prominently different from the other colours for all three measurements; It was expected to be less healthy, more flavour intense, and sweeter. Product popularity also influenced packaging liking. These results, however, did not translate to taste perception. Below we discuss in detail these findings and their potential mechanisms.

Regarding the effect of packaging colour on rated product expectations, the dark red package (RLH) was rated as less healthy than the blue and the light red packages. This is in line with research done by Huang and Lu (Huang & Lu, 2015, 2016), who found that utilitarian foods in blue packages are considered healthier than foods in red packages and Mai et al. (2016), who found that products in light packages are considered healthier than those in dark packages. A potential explanation is that these associations were created through learned symbolic meanings and stereotypes or previous exposures to healthy foods, which often have brighter and less saturated packaging colours compared to more unhealthy foods (Huang & Lu, 2016; Mai et al., 2016). Likewise, the red packages were rated higher on expected flavour intensity and sweetness than the blue packages. This is in line with the findings of Tijssen et al. (2017), who found that the red, dark and saturated package was rated highest on expected flavour intensity, and sweetness using the same package stimuli. Overall, the sensory results obtained in this study match the results by Mai et al. (2016), who found that there is an inverse association between the expected healthiness of a product and the expected flavour intensity. This association has been observed previously and is one of the difficulties companies face when trying to sell a healthy product (Mai et al., 2016), as healthy products may be perceived as less tasty and, as a consequence, might be less liked compared to their unhealthy counterparts (Raghunathan, Naylor, & Hoyer, 2006; Skaczkowski, Durkin, Kashima, & Wakefield, 2016).

On a physiological level, SCR increased while viewing red packages, and decreased for blue packages. No differences were found for HR. This is in line with Jacobs and Hustmyer (1974), who found that colour exposure did not change HR but changed SCR, with red eliciting a stronger SCR than blue. The difference in direction of colour-elicited SCR fits colour-related emotions. Clarke and Costall (2008) used a semi-structured interview to identify common descriptors of colours. They found that participants associated the colour red with the physically arousing emotions love and anger, whereas blue was described with less arousing emotions (soothing and comfortable). Therefore, we speculate that the calming effect of blue led to the decrease in SCR, whereas the arousing nature of red led to an increase. The ANS responses observed here support previous research suggesting that skin conductance activation may reflect emotional activation evoked through colour associations. On the other hand, HR is hypothesised to be discriminatory along the valence dimension (Beyts et al., 2017), implying that the different colours used here induced no difference in valence. Therefore, it can be hypothesised that small differences in packaging colour can evoke emotions that do not differ in their valence. No effects of product popularity on the ANS responses to the visualisation of the packages were found. This may indicate that a stronger personal relevance is necessary for this information to impact consumer responses at the ANS level.

The ratings of the product and ANS responses during the informed tasting were not influenced by package colour. However, the current research shows that popularity changed the SCR in such a way that when tasting yogurt low in popularity SCR decreased more than while tasting the highly popular yogurt. In previous research, a decrease in SCR and an unchanged or slightly positive HR, has been associated with the feeling of relief (Kreibig, 2010). This would match a potential anticipatory effect, which then switches into relief when the product with a low popularity score does not fulfil negative anticipations. Consumers tend to act according to descriptive norms especially when they feel uncertain about their choices (Huh, Vosgerau, & Morewedge, 2014). Therefore, it could be argued that participants did not expect the yogurt to be disappointing when it was labelled as popular. For the unpopular product, participants were probably unable to form positive expectations and thus felt relieved when the unpopular product was not disappointing. The current study was the first to examine the effect of descriptive norms on ANS responses and found an influence. The results presented here offer support for the hypothesis that popularity affects the expectations consumers may have about products. This may be associated with an increase in reward-related brain activity when the personal belief conforms with the group's belief (Robinson et al., 2014). Consumers value products popular amongst their peers more than those that are not popular (Mason, Dyer, & Norton, 2009), and have the tendency to conform to the group opinion to avoid discomfort (Klucharev, Hytonen, Rijpkema, Smidts, & Fernandez, 2009).

The results for the association between the explicit and implicit responses are ambiguous. While Wijk et al. (2014) found a positive association between HR and liking, Horio (2000) found a negative one. In this work, we found no association between the two. However, in line with Danner et al. (2014), a weak negative correlation was found between the liking of the taste and SCR. Furthermore, weak correlations between SCR and some sensory attributes were found. For all attributes the direction of the correlation differed between expectation and perception. For instance, the expected attributes had a positive correlation with SCR, whereas perceived flavour intensity was negatively correlated with SCR. This supports the assumption of a differential ANS activation depending on the type of sensory involvement (Wijk & Boesveldt, 2016).

Expected and perceived healthiness, as well as expected sweetness ratings, varied strongly between package colours and block position. The strong variation of perceived healthiness may be due to the nature of healthiness. This is based on credence, which in turn cannot be easily concluded from taste. Therefore, it is assumed that consumers infer healthiness from obtained package information, including the displayed colour (Mai et al., 2016). Based on the similar patterns between the expected and perceived healthiness of the yogurt, it is hypothesised that the taste perception may have assimilated towards the expectation previously elicited by the packaging. However, this hypothesis does not explain the fluctuations in the expectation ratings, which emphasises the importance of finding a reliable and interpretable method for combined sensory, consumer, and marketing research.

It is important to note that the measured differences in SCR were in the range of 0.0005–0.04 μS, which is below the range of SCR amplitudes typically found for non-food related stimuli (0.1–1.0 μS; Dawson, Schell, & Filion, 2007), but comparable to similar food related stimuli ($\leq$0.01 μS; Wijk et al., 2014). Therefore, the measured differences should be considered carefully until it can be confirmed that changes of this magnitude are meaningful. In the meantime, it may be assumed that SCR and ANS responses in general may not be suitable measures of food related expectations and perceptions. This is in line with Beyts et al. (2017) who compared the discriminatory power of self-reported liking, emotion ratings (using a consumer-led lexicon of ten emotion categories scored on VAS scales), the ANS responses HR and skin temperature, and facial electromyography. Their results show that emotion ratings discriminated best among beer aromas, while ANS responses discriminated least. This implies that other measures of emotional processes may be better suited to understanding consumer behaviour than ANS responses. Our results, along with the inconsistency of results found across the published articles in the field, indicate that ANS responses may not be sensitive and specific enough to assess subtle manipulations like food packaging and the effect of expectations on taste perception. One reason for this may be the fact that the influence of movement through sipping and chewing alone is larger than any of the influences detected as a response to the stimuli themselves (Wijk & Boesveldt, 2016). This implies that tests outside of strictly controlled laboratory sessions are difficult. This was confirmed by a pilot study conducted in a semi-controlled shopping environment that did not find any significant differences in SCR between a positive and a negative food package (Hurley et al., 2015). Based on our results and the literature we propose that a more cross-disciplinary approach using insights not only from sensory and consumer science, but also from the adjacent fields of psychology and physiology may yield new methods or combinations of existing methods for measuring implicit consumer responses. One potential method is the implicit association task, which has been found to measure underlying associations for example for food package colours (Piqueras-Fiszman & Spence, 2011; Tijssen et al., 2017) and food healthiness (Fulcher et al., 2016).

A strength of the current study is the adjustment for the consumed amount of yogurt per tasting. As previous studies mostly estimated consumption, it was unknown if these influences change ANS responses. Given that no influence of intake was found, it could be assumed that small differences in consumption sizes do not change ANS responses. This provides an avenue for future research, to assess ANS responses for higher intakes. A further strength of the study is the within subject design for the colour manipulation, leading to a more realistic setting. The study by Tijssen et al. (2017), which included the same stimuli, used a between subject design to study product expectations and sensory perceptions. They found that the colour manipulation was strong enough to change the sensory perception of a yogurt. However, while our results replicate the elicited expectations, we could not show an effect on sensory perception, indicating that the effect of elicited expectations on perception may not be as strong or reliable as previously assumed.

On the other hand, the study's results may have been partially influenced by participants guessing the study's aim, as indicated by the participants after the completion of the experiment. Although no significant differences in responses were found between the participants who guessed the aim and those that did not, an effect cannot be excluded. More specifically, it is expected that participants who guessed the aim may have been less involved in the study and therefore may have shown diminished ANS responses. Furthermore, it can be assumed that most people had preconceptions of the Optimel brand and potentially of the previously commercially available *'puur'* product. This may have influenced responses, as Optimel is a well-known brand of products in the Netherlands with 0% fat and no added sugar catering to the light product oriented consumer (Optimel Eurosparen, 2017). This influence may explain the 10% decrease in liking between the blindly sampled yogurt during the practice run and the following experimental runs (data not shown). This emphasises the fact that extrinsic food cues play an important role in product perception and need to be included in consumer tests, which in turn need to be able to pick up on these effects. Additionally, an international study population was used, whereas the stimulus product was Dutch. This may have influenced results by leading to more diverse responses through a difference in understanding. In order to adjust for this, nationality (Dutch and Belgian vs. other) was included in the statistical models. Results showed that nationality only had a significant influence on perceived sweetness and HR while viewing the package image. Furthermore, it is known that the frame of reference, as well as colour associations are different between countries (Piqueras-Fiszman & Spence, 2011). Since all participants were living in the Netherlands at the time of the study and were therefore confronted with Dutch products on a daily basis it was assumed that this effect was diminished. Nevertheless, it would be of interest to conduct a study with a more homogeneous study population to see if the results can be confirmed.

5. Conclusion

This study was the first to examine explicit and ANS responses to manipulations of package colour and the descriptive social norm product popularity. We found that expectations of the product differed between package images but did not influence taste perception. Both explicit ratings and SCR supported this result, however, no predictive associations could be found between these measures. Furthermore, product popularity was shown to lead to a difference in package liking as well as SCR to the taste of yogurt indicating that social descriptive norms may play a role in modifying expectations. However, contrary to what seems to be assumed, the current results show that ANS responses may not be sensitive enough to detect differences in cognitive processes elicited by packaging cues.

Acknowledgements

The authors would like to thank Irene Tijssen for her support with the study design. RPGG was supported by a grant from the Netherlands Organization for Scientific Research, FrieslandCampina, Amersfoort, the Netherlands and Unilever R & D Vlaardingen (NL) (FCBG13-03). PAMS is supported by the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement 607310 (Nudge-It).

References

Ares, G., & Deliza, R. (2010). Studying the influence of package shape and colour on consumer expectations of milk desserts using word association and conjoint analysis. *Food Quality and Preference*, *21*(8), 930–937. http://dx.doi.org/10.1016/j.foodqual.2010.03.006.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823.

Bearden, W. O., & Rose, R. L. (1990). Attention to social comparison information: An individual difference factor affecting consumer conformity. *Journal of Consumer Research*, *16*(4), 461–471. http://dx.doi.org/10.1086/209231.

Beyts, C., Chaya, C., Dehrmann, F., James, S., Smart, K., & Hort, J. (2017). A comparison of self-reported emotional and implicit responses to aromas in beer. *Food Quality and Preference*, *59*, 68–80. http://dx.doi.org/10.1016/j.foodqual.2017.02.006.

Blundell, J., de Graaf, C., Hulshof, T., Jebb, S., Livingstone, B., Lluch, A., & Westerterp, M. (2010). Appetite control: methodological aspects of the evaluation of foods. *Obesity Reviews*, *11*(3), 251–270. http://dx.doi.org/10.1111/j.1467-789X.2010.00714.x.

Bosshard, S. S., Bourke, J. D., Kunaharan, S., Koller, M., Walla, P., & Heinonen, J. (2016). Established liked versus disliked brands: Brain activity, implicit associations and explicit responses. *Cogent Psychology*, *3*(1), 347. http://dx.doi.org/10.1080/23311908.2016.1176691.

Clarke, T., & Costall, A. (2008). The emotional connotations of color: A qualitative investigation. *Color Research & Application*, *33*(5), 406–410. http://dx.doi.org/10.1002/col.20435.

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. http://dx.doi.org/10.1016/j.foodres.2014.06.003.

Davidenko, O., Delarue, J., Marsset-Baglieri, A., Fromentin, G., Tome, D., Nadkarni, N., & Darcel, N. (2015). Assimilation and contrast are on the same scale of food anticipated-experienced pleasure divergence. *Appetite*, *90*, 160–167. http://dx.doi.org/10.1016/j.appet.2015.03.006.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. *Handbook of Psychophysiology*, *2*, 200–223.

Fulcher, E., Dean, A., & Trufil, G. (2016). Neurosense and packaging: Understanding consumer evaluations using implicit technology. In P. Burgess (Ed.), *Woodhead Publishing Series in Food Science, Technology and Nutrition. Integrating the Packaging and Product Experience in Food and Beverages*(pp. 121–138). Woodhead Publishing. http://dx.doi.org/10.1016/B978-0-08-100356-5.00006-1.

Higgs, S. (2015). Social norms and their influence on eating behaviours. *Appetite*, *86*, 38–44. http://dx.doi.org/10.1016/j.appet.2014.10.021.

Horio, T. (2000). Effects of various taste stimuli on heart rate in humans. *Chemical Senses*, *25*(2), 149–153. http://dx.doi.org/10.1093/chemse/25.2.149.

Huang, L., & Lu, J. (2015). Eat with your eyes: Package color influences the expectation of food taste and healthiness moderated by external eating. *Marketing Management*, *25*(2), 71–87.

Huang, L., & Lu, J. (2016). The impact of package color and the nutrition content Labels on the perception of food healthiness and purchase intention. *Journal of Food Products Marketing*, *22*(2), 191–218. http://dx.doi.org/10.1080/10454446.2014.1000434.

Huh, Y. E., Vosgerau, J., & Morewedge, C. K. (2014). Social defaults: Observed choices become choice defaults. *Journal of Consumer Research*, *41*(3), 746–760. http://dx.doi.org/10.1086/677315.

Hurley, R. A., Hutcherson, D. E., Tonkin, C. E., Dailey, S. B., & Rice, J. C. (2015). Measuring physiological arousal towards packaging: tracking electrodermal activity within the consumer shopping environment. *Journal of Applied Packaging Research*, *7*(3).

Jacobs, K. W., & Hustmyer, F. E., Jr (1974). Effects of four psychological primary colors on GSR, heart rate and respiration rate. *Perceptual and Motor Skills*, *38*(3), 763–766. http://dx.doi.org/10.2466/pms.1974.38.3.763.

Klucharev, V., Hytonen, K., Rijpkema, M., Smidts, A., & Fernandez, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron*, *61*(1), 140–151. http://dx.doi.org/10.1016/j.neuron.2008.11.027.

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: a review. *Biological Psychology*, *84*(3), 394–421. http://dx.doi.org/10.1016/j.biopsycho.2010.03.010.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest: Tests in Linear Mixed Effects Models. Available from https://CRAN.R-project.org/package=lmerTest.

Lenth, R. V. (2016). Least-squares means: The R package ismeans. *Journal of Statistical Software*, *69*(1), http://dx.doi.org/10.18637/jss.v069.i01.

Mai, R., Symmank, C., & Seeberg-Elverfeldt, B. (2016). Light and pale colors in food packaging: When does this package cue signal superior healthiness or inferior tastiness? *Journal of Retailing*, *92*(4), 426–444. http://dx.doi.org/10.1016/j.jretai.2016.08.002.

Martin, G. (1998). Human electroencephalographic (EEG) response to olfactory stimulation: Two experiments using the aroma of food. *International Journal of Psychophysiology*, *30*(3), 287–302. http://dx.doi.org/10.1016/S0167-8760(98)00025-7.

Mason, M. F., Dyer, R., & Norton, M. I. (2009). Neural mechanisms of social influence. *Organizational Behavior and Human Decision Processes*, *110*(2), 152–159. http://dx.doi.org/10.1016/j.obhdp.2009.04.001.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. http://dx.doi.org/10.3758/s13428-011-0168-7.

Melnyk, V., van Herpen, E., & Trijp, H. (2010). The influence of social norms in consumer decision making: A meta-analysis. NA-Advances in Consumer Research Volume 37.

Mojet, J., Dürrschmid, K., Danner, L., Jöchl, M., Heiniö, R.-L., Holthuysen, N., & Köster, E. (2015). Are implicit emotion measurements evoked by food unrelated to liking? *Food Research International*, *76*, 224–232. http://dx.doi.org/10.1016/j.foodres.2015.06.031.

Nederkoorn, C., Smulders, F. T. Y., & Jansen, A. (2000). Cephalic phase responses, craving and food intake in normal subjects. *Appetite*, *35*(1), 45–55. http://dx.doi.org/10.1006/appe.2000.0328.

Nook, E. C., & Zaki, J. (2015). Social norms shift behavioral and neural responses to foods. *Journal of Cognitive Neuroscience*, *27*(7), 1412–1426. http://dx.doi.org/10.1162/jocn_a_00795.

Ohla, K., Toepel, U., Le Coutre, J., & Hudry, J. (2012). Visual-gustatory interaction: Orbitofrontal and insular cortices mediate the effect of high-calorie visual food cues on taste pleasantness. *PLoS One*, *7*(3), e32434. http://dx.doi.org/10.1371/journal.pone.0032434.

Okamoto, M., & Dan, I. (2013). Extrinsic information influences taste and flavor perception: A review from psychological and neuroimaging perspectives. *Seminars in Cell & Developmental Biology*, *24*(3), 247–255. http://dx.doi.org/10.1016/j.semcdb.2012.11.001.

Optimel Eurosparen (2017). Over Optimel. Available from https://www.optimel.nl/producten/over-optimel.

Piqueras-Fiszman, B., & Spence, C. (2011). Crossmodal correspondences in product packaging. Assessing color-flavor correspondences for potato chips (crisps). *Appetite*, *57*(3), 753–757. http://dx.doi.org/10.1016/j.appet.2011.07.012.

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*(Part A), 165–179. http://dx.doi.org/10.1016/j.foodqual.2014.09.013.

Piqueras-Fiszman, B., Velasco, C., Salgado-Montejo, A., & Spence, C. (2013). Using combined eye tracking and word association in order to assess novel packaging solutions: A case study involving jam jars. *Food Quality and Preference*, *28*(1), 328–338. http://dx.doi.org/10.1016/j.foodqual.2012.10.006.

R Core Team. (2013). R: A language and environment for statistical computing. Available from https://www.R-project.org/.

Raghunathan, R., Naylor, R. W., & Hoyer, W. D. (2006). The unhealthy = tasty intuition and its effects on taste inferences, enjoyment, and choice of food products. *Journal of Marketing*, *70*(4), 170–184. http://dx.doi.org/10.1509/jmkg.70.4.170.

Robinson, E. (2015). Perceived social norms and eating behaviour: An evaluation of studies and future directions. *Physiology & Behavior*, *152*(Part B), 397–401. http://dx.doi.org/10.1016/j.physbeh.2015.06.010.

Robinson, E., Fleming, A., & Higgs, S. (2014). Prompting healthier eating: Testing the use of health and social norm based messages. *Health Psychology*, *33*(9), 1057–1064. http://dx.doi.org/10.1037/a0034213.

Roininen, K., Tuorila, H., Zandstra, E. H., de Graaf, C., Vehkalahti, K., Stubenitsky, K., & Mela, D. J. (2001). Differences in health and taste attitudes and reported behaviour among Finnish, Dutch and British consumers: a cross-national validation of the Health and Taste Attitude Scales (HTAS). *Appetite*, *37*(1), 33–45. http://dx.doi.org/10.1006/appe.2001.0414.

Rousmans, S., Robin, O., Dittmar, A., & Vernet-Maury, E. (2000). Autonomic nervous system responses associated with primary tastes. *Chemical Senses*, *25*(6), 709–718.

http://dx.doi.org/10.1093/chemse/25.6.709.
Skaczkowski, G., Durkin, S., Kashima, Y., & Wakefield, M. (2016). The effect of packaging, branding and labeling on the experience of unhealthy food and drink: A review. *Appetite*, *99*, 219–234. http://dx.doi.org/10.1016/j.appet.2016.01.022.
Spinelli, S., & Niedziela, M. (2016). Emotion measurements and application to product and packaging development. *Integrating the Packaging and Product Experience in Food and Beverages: A Road-Map to Consumer Satisfaction*, 77–119.
Stok, F. M., Ridder, D. T. D.de, Vet, E.de, & Wit, J. B. F.de (2014). Don't tell me what I should do, but what others do: the influence of descriptive and injunctive peer norms on fruit consumption in adolescents. *British Journal of Health Psychology*, *19*(1), 52–64. http://dx.doi.org/10.1111/bjhp.12030.
Surowiecki, J. (2005). *The wisdom of crowds* (1st ed.). New York, NY: Anchor Books.
Tijssen, I., Zandstra, E. H., Graaf, C.de, & Jager, G. (2017). Why a 'light' product package should not be light blue: Effects of package colour on perceived healthiness and attractiveness of sugar- and fat-reduced products. *Food Quality and Preference*, *59*, 46–58. http://dx.doi.org/10.1016/j.foodqual.2017.01.019.
van Strien, T., Frijters, J. E. R., Bergers, G. P. A., & Defares, P. B. (1986). The dutch eating behavior questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. *International Journal of Eating Disorders*, *5*(2), 295–315. http://dx.doi.org/10.1002/1098-108x(198602)5:2<295::aid-eat2260050209>3.0.co;2-t.
van Trijp, H. C., & Schifferstein, H. N. J. (1995). Sensory analysis in marketing practice: comparison and integration. *Journal of Sensory Studies*, *10*(2), 127–147. http://dx.doi.org/10.1111/j.1745-459X.1995.tb00010.x.
Wijk, R. A.de, & Boesveldt, S. (2016). Responses of the autonomic nervous system to flavors. *Multisensory Flavor Perception*(pp. 249–268). Woodhead Publishing. http://dx.doi.org/10.1016/B978-0-08-100350-3.00013-4.
Wijk, R. A.de, He, W., Mensink, M. G. J., Verhoeven, R. H. G., & Graaf, C.de (2014). ANS responses and facial expressions differentiate between the taste of commercial breakfast drinks. *PLoS One*, *9*(4), http://dx.doi.org/10.1371/journal.pone.0093823.
Wijk, R. A.de, Kooijman, V., Verhoeven, R. H. G., Holthuysen, N. T. E., & Graaf, C.de (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. http://dx.doi.org/10.1016/j.foodqual.2012.04.015.
Willemsen, G. H. M., de Geus, E. J. C., Klaver, C. H. A. M., van Doornen, L. J. P., & Carroll, D. (1996). Ambulatory monitoring of the impedance cardiogram. *Psychophysiology*, *33*(2), 184–193. http://dx.doi.org/10.1111/j.1469-8986.1996.tb02122.x.
Zielinski, P. (2016). An arousal effect of colors saturation: A study of self-reported ratings and electrodermal responses. *Journal of Psychophysiology*, *30*(1), 9–16. http://dx.doi.org/10.1027/0269-8803/a000149.