Food Quality and Preference 48 (2016) 274–282

Contents lists available at ScienceDirect

Food Quality and Preference

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

Could Time–Intensity by a trained panel be replaced with a progressive profile done by consumers? A case on chewing-gum

M.V. Galmarini [a,b,d,*], R. Symoneaux [c], M. Visalli [d], M.C. Zamora [a,b], P. Schlich [d]

[a] Facultad de Ciencias Agrarias, Pontificia Universidad Católica Argentina, Buenos Aires, Argentina
[b] CONICET, Buenos Aires, Argentina
[c] Groupe ESA, UPSP GRAPPE, SFR QUASAV 4207, Angers, France
[d] INRA, UMR1324 Centre des Sciences du Goût et de l'Alimentation, Dijon, France

ARTICLE INFO

Article history:
Received 7 July 2015
Received in revised form 20 October 2015
Accepted 26 October 2015
Available online 30 October 2015

Keywords:
In-home test
Dynamic methods
Progressive profile
Temporal description
Consumers

ABSTRACT

How to evaluate a chewing-gum profile in a reliable cost and time-efficient manner giving the industry the insight they need on their new products? The aim of the present work was to compare the temporary descriptive results obtained by a reference method such as Time–Intensity (T–I) done by a trained panel to those acquired by a progressive profile (PP) done by regular consumers in *in-home* conditions. The evolution of four different attributes (sweetness, mint aroma, hardness and freshness) during time was studied by each method. Results were compared on the basis of three different parameters: the maximum intensity reached (I_{max}), the time to reach this maximum intensity (T_{Imax},); and the area under the curve (AUC), which integrated both time and intensity. Sample discrimination was good for the trained panel and for the consumers. Comparable results were obtained for the parameter AUC for all attributes, showing a similar global description of all samples by both methods and groups. However, differences were found in the T_{Imax}. According to the obtained results, T–I still gives more detailed information and should not be replaced when small changes are studied. However, if looking to validate the sensory description of a different new prototype, the PP done by consumers in *in-home* conditions might be a very interesting option being more cost and time efficient.

1. Introduction

Constant product innovation in the chewing gum sector can perk up sales opportunities (Research-and-Markets, 2014). Innovation might include sweetener changes, new flavors, center-filled chewing gums, different textures; all characteristics for which the sensory perception changes over time (Lenzi et al., 2012; Piggott, 2000; Saint-Eve et al., 2011; Song, Knutsen, Broderick, & Seielstad, 2010). Therefore their description requires dynamic sensory methodologies able to give reliable information on the evolution of product perception along time. In addition, techniques should be cost and time-efficient in order to give the industry the insight they need on their new products.

Time–Intensity (T–I) curves were one of the first continuous ways of registering changes in the perception of a certain attribute (Cliff & Heymann, 1993; Lee & Pangborn, 1986). This method has been highly studied and its data analysis discussed and improved (Eilers & Dijksterhuis, 2004; Garrido, Calviño, & Hough, 2001; McGowan & Lee, 2006; Ovejero-López, Bro, & Bredie, 2005; Piggott, 2000), being one of the reference methods in terms of dynamic description. However, it has never been broadly used as a routine method mainly for being time-consuming and expensive. It is time consuming because it requires a high level of training of the assessors, and it is also limited to studying one attribute at a time (Labbe, Schlich, Pineau, Gilbert, & Martin, 2009; Pineau et al., 2009). Different attributes can be measured in repeated T–I tests to obtain composite T–I profiles (Devezeaux de la vergne, Van delft, Van develde, Van boekel, & Stieger, 2015), but this means that the amount of sessions increases rapidly. To cope with the mentioned disadvantages, Duizer, Bloom, and Findlay (1997) developed the Dual Attribute Time–Intensity (DATI). Working with chewing-gum, they measured two attributes at the same time by using a horizontal and a vertical scales. They demonstrated that measuring two sensory attributes at the same time in a continuous manner was possible, and allowed the sensory scientists to determine the interaction of perceptions in the mouth during mastication. However, this only allowed quantifying two attributes per session, still required a high level of training and it rose the

* Corresponding author at: Facultad de Ciencias Agrarias, UCA. Cap. Gral Ramon Freire 183, 1429, Buenos Aires, Argentina.
E-mail address: mgalmarini@gmail.com (M.V. Galmarini).

http://dx.doi.org/10.1016/j.foodqual.2015.10.006
0950-3293/

question of how many tasks can a panelist attend to over a period of time in a simultaneous and continuous manner.

In order to measure more attributes at a time, Temporal Dominance of Sensations (TDS) (Labbe et al., 2009; Pineau et al., 2009) offers an advantage. But this technique focuses more on the relative effect of attributes (dominance) rather than on the intensity of each individual sensory characteristic over time. Most recently, Kuesten, Bi, and Feng (2013) came up with the Multi-Attribute Time–Intensity (MATI) approach which allowed for multiple attributes to be collected intermittently using a paced rate through repeated cycles (one sec apart). They explored their technique with taffy as a product, having a total evaluation time of 2.5 min and found that even though multiple attributes were evaluated simultaneously, it required a high level of training of the panelists.

Progressive Profiling, first described by Jack, Piggott, and Paterson (1994) is another dynamic technique which uses discrete instead of continuous time intervals (Jack et al., 1994; Piggott, 2000). A series of line scales can be presented on the same screen (practically a maximum of 5, depending on the product) and assessors repeatedly profile the sample during tasting at fixed time intervals. The panelist quantifies each attribute in the same way as in a standard descriptive profiling. The time assessors have to answer and the time in-between answers depends on the product and the used protocol. This technique has the benefit of being a simple descriptive (qualitative and quantitative) analysis where attributes are fixed, making their measurement sequential and repetitive, reducing the cognitive load and needing less training than T–I measurements (Devezeaux de la vergne et al., 2015). Compared to the aforementioned techniques, it presents the disadvantage of being a discontinuous way of data collection, which results in fewer data points for modeling the curves as a function of time. However, in a product such as chewing-gum which has a long consumption time span (minutes) in comparison to most food products (just several sec per bite), this could even be an advantage trying to avoid the fatigue which can arise from continuously measuring an attribute over long periods of time (e.g. 10 min). Moreover, the lower training needed to perform the task raises the question if it could be performed without previous training. However, very few references can be found comparing the Progressive Profiling to other temporal methods (Devezeaux de la vergne et al., 2015).

In their review on novel methods for product characterization, Varela and Ares (2012) pointed out that in certain cases the training period for developing a Quantitative Descriptive Analysis (QDA) could be omitted and that the sensory characterization could be performed by consumers (Husson, le Dien, & Pagès, 2001; Worch, Lê, & Punter, 2010). In this approach, consumers are asked to rate the intensity of a fixed set of sensory attributes using intensity scales, as it is commonly done with trained assessors in descriptive analysis, but descriptors are provided to consumers with no further training in attribute recognition or quantification. Even if this approach has been traditionally not recommended (Lawless & Heymann, 1998; Stone & Sidel, 2003), studies have reported that results from sensory characterization performed by 50–100 consumers with intensity scales are similar to those provided by trained assessor panels (Husson et al., 2001; Worch et al., 2010). They showed that the apparent lack of consensus and high variability in consumer responses could be compensated by a large sample size ($n \geqslant 50$), making the use of consumers for a descriptive task viable.

Nonetheless, handling this number of consumers in the sensory lab, particularly for long tests, would mean an increase in costs and a high reduction of time availability. For this reason, in the present work the profiling task done by consumers was performed in *in-home* conditions with the help of a web application. In addition, an in-home sensory evaluation enables to collect sensory data in more real conditions (Martin, Visalli, Lange, Schlich, & Issanchou, 2014) and closer to consumers' habits. In this way, the consumer test becomes cheaper and more efficient since the number of consumers performing the test is independent from the facilities of the sensory lab (Galmarini, Symoneaux, Visalli, Zamora, & Schlich, 2014) while consumers can evaluate the product in a natural environment.

It was the aim of the present work to compare the descriptive results obtained by a reference method to those acquired in a less traditional but more time-efficient fashion. In this way, the T–I curves done on mint chewing-gums by a trained panel evaluating one attribute at a time in a sensory laboratory were compared to a progressive profile done by regular consumers in *in-home* conditions; attempting to find the most suitable method for a dynamic description of such product.

2. Materials and methods

2.1. Samples

Three different commercial Argentinean chewing gums (hereon CH-1, CH-2, CH-3) were used for this study. The three were mint flavoured, not sugar coated and sugar free. They were all intended for the same market segment: young consumers (16–30 years old) of medium/upper social classes, and the price per unit of chewing gum was of around AR$ 0.85.

2.2. Time–Intensity (T–I) measurements

2.2.1. Trained panel

Nine trained assessors, three men and six women (aged 18–24 years-old) were recruited for the Time–Intensity (T–I) study. They were part of the permanent panel of the Sensory Laboratory of the Universidad Católica Argentina and were paid for participating in this study. All of them had previous experience on T–I methodology.

Assessors followed four training sessions (30 min long each) in order to differentiate attributes, become acquainted with the samples and with the T–I software.

During the first session, assessors were presented the attributes to be measured: sweetness, mint aroma, freshness and hardness. For each attribute, references were given as follows: sweetness, intensive sweetener (Hileret, Argentina; aspartame and acesulfame-K); mint, mint aroma provided by Givaudan (Argentina); freshness, cooling agent (containing as the active ingredient acyclic tertiary and secondary carboxamides, Givaudan, Argentina); and non-flavoured chewing-gum for hardness. Three concentrations were presented in duplicate for each reference (except for hardness): sweetness: 0.2, 0.3 and 0.4% w/v; mint aroma: 0.2, 0.4, 0.6% w/w; cooling agent: 0.03, 0.05, 0.07% w/v. Assessors were asked to identify the references and place them on a 10 cm unstructured scale (one scale per attribute).

Throughout the second and third sessions, panelists became familiar with the samples to be tested. They evaluated them in a static manner: they quantified the 4 given attributes on unstructured scales after 30 s of chewing.

The last training session was devoted to the use of the T–I methodology in order to coordinate sample intake and constant data recording. This was done evaluating the sweet descriptor on a different chewing gum which was not to be characterized afterwards.

2.2.2. Testing protocol

The evolution of the four different attributes (sweetness, mint aroma, hardness and freshness) during time was studied according to ISO TC 34/SC 12 N 385 (ISO, 1999) using a computer software specially designed for this purpose. Assessors used a mouse to move a cursor along a line that represented a 100 mm unstructured line scale on the monitor. For each measurement, data was automatically recorded every 0.35 s. The software provided the T–I curve as well as eight parameters which described it: maximum intensity reached (I_{max}), time elapsed to maximum intensity (T_{max}), total duration of perceived sensation (T_{dur}), time for intensity to decline to half its maximum value (T_{50max}), area under curve (AUC), rate of increase of perception (R_{inc}), rate of decrease of perception (R_{dec}) and plateau time (time during which perceived intensity remained constant; T_{plat}).

During each session, panelists evaluated all three samples (presented in a random order) for one attribute. A total of 8 measurement sessions were done in order to evaluate each sample, in duplicate for the 4 attributes. The first two sessions of measurement were devoted to sweetness, the next for hardness, then for mint and the final two for freshness. Sessions lasted between 45 and 60 min (depending on the attribute measured) and a 10 min break between samples was imposed to avoid sensory fatigue. This resulted in approximately 8 h of work per panelist (plus training) and a total of 216 curves to be analyzed.

2.3. Progressive Profiling (PP)

2.3.1. Consumers

A total of 50 regular mint chewing-gum consumers with no previous experience in sensory evaluation participated in the test, 54% females and 46% males, ages ranging from 19 to 32 years old. They were recruited among students and staff of the Universidad Católica Agentina (Buenos Aires), based on their willingness to participate and their frequency of consumption of mint chewing-gum. Of this population, 67% consumed mint chewing gum at least 2–3 times a week and the rest consumed 2–3 times a month.

2.3.2. Testing protocol

The 50 consumers carried out the test from their homes and data was acquired by TimeSens online software (www.timesens.com). Samples were given out personally at the Sensory Laboratory at the Universidad Católica Argentina, in a sealed envelope containing all three samples and the instructions needed to access their online session. At the same time, they were explained that the test could be done at any moment of the day, evaluating only one sample a day, needing a computer or a tablet with Internet service. They were also instructed that what they had to do was an objective measurement, and that they did not have to base their answers on their liking of the sample.

The online session began by explaining the way to carry out the test (Fig. 1a) and the sample to be tested was instructed (Fig. 1b). The order of presentation of samples was randomized among consumers and access to the web site was limited ensure that consumers only evaluated one sample a day.

Intensity of the four studied attributes (sweetness, hardness, mint aroma and freshness) was evaluated at intervals of 45 s on a VAS scale along a 10 min period (Fig. 1c). They were asked to rate the four attributes at the same time, and they had 15 s to provide their answer. Assessors were proposed small readings as background task in-between evaluations (Fig. 1d; Galmarini et al., 2014). The main aim of this background task was to avoid boredom but yet allow a level of concentration on the sensory task along the tasting period. This reading was in the form of curious facts (“Did you know…?”) which changed every 15 s.

The previous ratings were presented to consumers from one scoring moment to another (except for the first evaluation) and they were allowed to change it or to leave it the same in relation to what they perceived. In this way, 11 intensity scores were obtained per assessor for each chewing gum for the 10 min consumption period, without any missing data. Since consumers evaluated each chewing gum only once, the total task required only 30 min net of work. The order in which the attributes were presented was the same all along the test for each consumer but it was randomized among them in order to avoid the halo effect (Lawless & Heymann, 1999).

2.4. Data analysis

2.4.1. T–I data

For each of the T–I curves the parameters I_{max}, T_{Imax}, T_{dur}, T_{50max}, AUC, R_{inc}, R_{dec}, T_{plat} (described in Section 2.2.2) were calculated for each attribute. These were then analyzed by a two-way Analysis of Variance (ANOVA) with Subject as random effect and posterior Least-Significant Difference (LSD) tests.

2.4.2. PP data

Discrete PP data points were used for recreating curves of intensity as a function of time. After this continuous representation of the data, similar parameters as for T–I were obtained: AUC, T_{Imax}, I_{max}, R_{dec} and R_{inc}. These were also analyzed by ANOVA (subject as random effect and posterior Least-Significant Difference (LSD) tests). The parameters T_{dur}, T_{50max} and T_{plat} were not calculated since the average intensity of these attributes did not reach zero (or half the maximum intensity) after 10 min.

2.4.3. Comparison of both methodologies: T–I vs. PP

Results obtained by the two methodologies were compared by three complementary methods: (1) by visual assessment of the obtained Intensity vs. Time curves; (2) by MANOVA and further Canonical Variance Analysis (CVA) (Peltier, Visalli, & Schlich, 2015) representation of three parameters used to summarize information from the curves; (3) by comparing the trajectory map at different time-points. All three methods are further described below.

For the purpose of comparison, average Intensity vs. Time curves were produced by attribute, for each sample on both data sets (T–I and PP). This is a visual method which enables comparison but which does not allow concluding since no statistical test is performed on this curves.

Therefore, in order to be able to better compare results obtained by both methods, three different parameters were chosen: a time dependent parameter, T_{Imax} (even though it was not significantly different among samples for all parameters); an intensity related parameter, I_{max} and a parameter which integrates both time and intensity, AUC. With these parameters for each attribute (averaged over replicates in the case of T–I), two independent MANOVA represented by CVA (Peltier et al., 2015) were performed on the data set from each panel. This was done using TimeSens software.

Finally, trajectory maps of both dynamic methods were compared. In order to obtain this map, a Principal Component Analysis (PCA) was done considering the four attributes (mint, freshness, hardness and sweetness) as variables and using 11 time points. This choice was made based on the 11 discrete recordings obtained by PP, as for T–I measurements, only the values obtained at this particular moments were taken into consideration for this analysis (average value among panelists at each time point). The data table submitted to a covariance PCA was thus composed of 66 observations (3 products in 2 methods by 11 points) and 4 variables (sensory attributes). Data was centered for each method before doing the PCA. These analyses were performed with the R software (version 3.1.2, R Core Team) and FactoMineR package (Husson, Josse, Le, & Mazet, 2013).

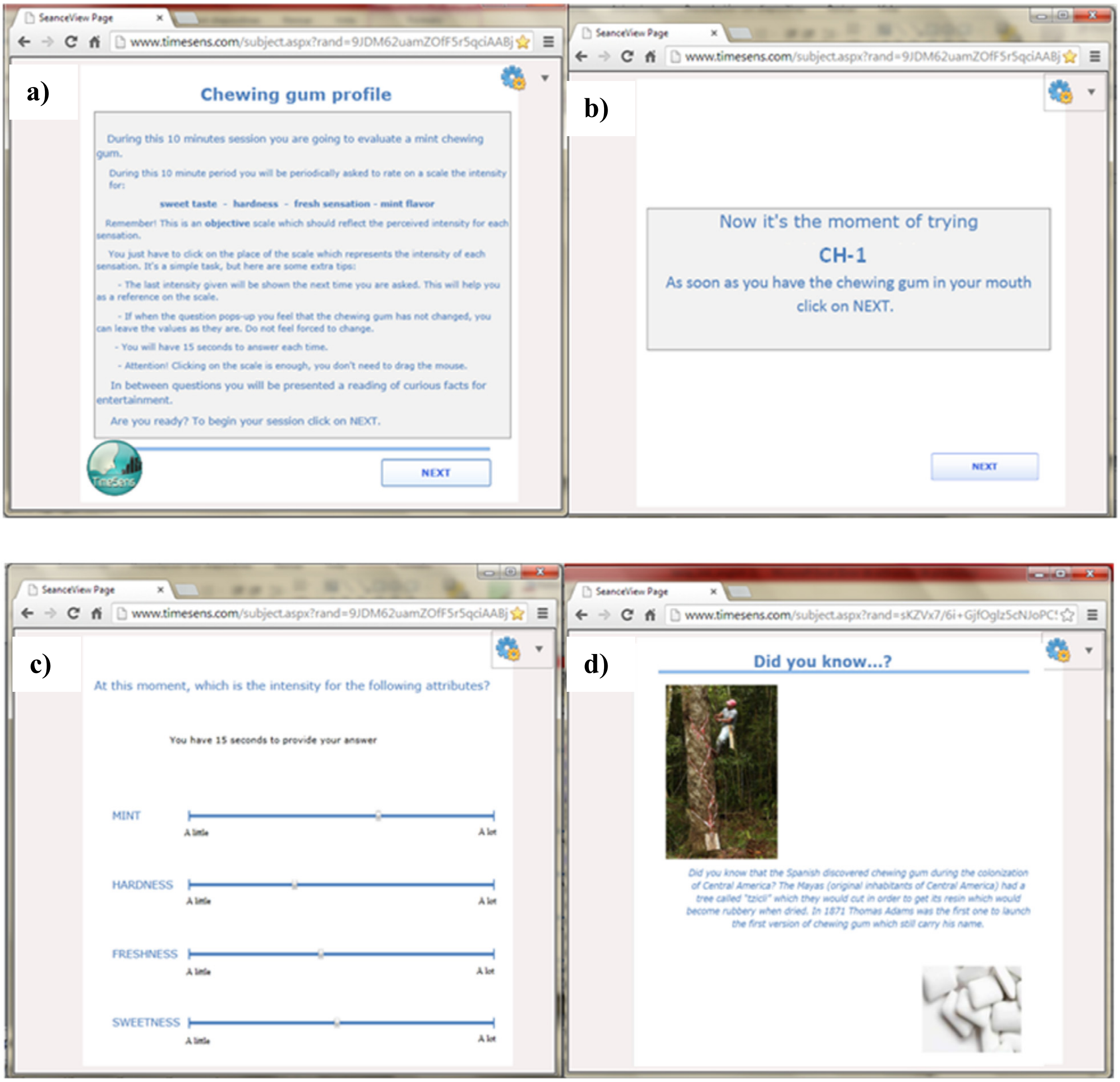

Fig. 1. (a–d) Protocol of data acquisition in *in-home* conditions. (a) General instructions given to consumers before the test; (b) instructions given to consumers before trying a particular sample; (c) evaluation screen. These appeared every 45 s, after the first time, the precedent given notes were presented to as a reference; (d) example of the reading proposed as a background task. Every 22.5 s the "Did you know…?" changed. In all cases, this information was presented in Spanish; it was here translated to English from the original screen for the purpose of clarity.

3. Results

Table 1 presents the results obtained for the analysis of the eight parameters of the T–I data from the trained panel. It can be observed that significant differences among samples were found in every attribute for most parameters; AUC and I_{max} were always highly significant ($p < 0.001$). In particular for mint, samples were also different in terms of T_{dur} and R_{inc}. For sweetness there were also differences in the R_{dec} as well as in T_{dur}. For freshness and hardness the R_{dec} was also different among samples. Finally, they were different for T_{Imax} in terms of freshness and for T_{pla} in the case of hardness.

In Table 2 the results of the ANOVA performed on the five parameters of the PP are presented. In terms of sample discrimination, the parameters I_{max} and AUC were significantly different among samples for the four attributes ($p < 0.05$–0.001; as well as for the T–I results). Only in the case of hardness, the parameters T_{Imax} and R_{dec} were also significantly different among samples. For freshness, the R_{inc} was the parameter significantly different in addition to AUC and I_{max}. It was also observed that differences among consumers (as opposed to differences among panelists) were significant for several parameters (*F*-values not shown), denoting more heterogeneity among them in the use of scale, certainly due to a lack of training on the attributes.

As a first approach for method evaluation, the complete average curves of intensity vs. time of each method (by attribute) are presented in Fig. 2a through h. For the purpose of comparison, the scale used by consumers was multiplied by a factor of 10 in order to have both 0–100 scales.

Table 1
ANOVA results for Time–Intensity parameters.

Attribute	Sample *F*-values	Attribute	Sample *F*-values
Mint I_{max}	52.39***	Sweetness I_{max}	30.98***
Mint Ti_{max}	7.84*	Sweetness Ti_{max}	0.03
Mint AUC	31.16***	Sweetness AUC	10.10**
Mint T_{dur}	10.24***	Sweetness T_{dur}	92.20***
Mint T_{50max}	5.89**	Sweetness T_{50max}	3.74*
Mint R_{dec}	0.57	Sweetness R_{dec}	25.99***
Mint R_{inc}	4.27*	Sweetness R_{inc}	0.31
Mint T_{plat}	3.21	Sweetness T_{plat}	1.03
Freshness I_{max}	45.82***	Hardness I_{max}	42.16***
Freshness Ti_{max}	13.68***	Hardness Ti_{max}	1.22
Freshness AUC	123.03***	Hardness AUC	16.72***
Freshness T_{dur}	[a]	Hardness T_{dur}	[a]
Freshness T_{50max}	3.43	Hardness T_{50max}	2.88
Freshness R_{dec}	12.89***	Hardness R_{dec}	51.49***
Freshness R_{inc}	1.52	Hardness R_{inc}	3.13
Freshness T_{plat}	1.46	Hardness T_{plat}	17.36***

[a] The perception of the sensation continued after 10 min of test.
*** $p < 0.001$.
** $p < 0.01$.
* $p < 0.05$.

For the attribute sweetness (Fig. 2a and b), in both groups the average values for I_{max} ranged between 45 and 65, showing that consumers were well able to quantify this basic taste on a scale, even without previous training. Both groups agreed on which sample was the least sweet (CH-1) giving to it nearly the same maximum intensity. However, there were differences in terms of the lasting of the sensation. While for the trained panel the sweet taste was extinguished after around 6 min in samples CH-1 and CH-3, in average consumers perceived this taste during the 10 min of test for the three samples.

As for hardness (Fig. 2c and d) a very similar temporal profile can be observed (sample ranking and use of the scale) between the two groups. Sample CH-3 was the least hard and the one that became softer with time. As for the other two samples, they were harder and more stable along time according to the two different panels.

In the description of freshness and mint (Fig. 2e–h respectively) there was a clear difference in the use of scales between the two groups and in the lasting for the case of mint. Consumers gave higher intensity values and also a longer lasting time to the perception of mint. These differences will be further analyzed in the discussion section.

Looking at the curves obtained by both groups, rich information can be obtained. However, no statistical test is performed on data. To further evaluate differences between methods (and panels), using the three main parameters which describe the above presented curves (see materials and methods section), two independent CVA (Peltier et al., 2015) representations of the MANOVA's tests were done for each panel (Fig. 3a and b). In this way sample discrimination and description by I–T done by the trained panel (Fig. 3a) and the PP done by consumers (Fig. 3b) were compared. In these graphs the size of the ellipses is a reflection of panel agreement (variability of subjects scores around the mean, Peltier et al., 2015). This is why ellipses were smaller for the trained panel (even if there were more consumers in the untrained panel), showing a higher agreement in terms of sample description. However, discrimination among samples was good in both panels since ellipses did not overlap in any case. Nonetheless, discrimination was bigger by the trained panel, as reflected by the MANOVA statistics which is twice as large as that of consumers ($F = 16.9$ for the trained panel and 7.3 for consumers).

Finally, in order to see the complete evolution (considering the change in the 4 attributes at the same time) of each sample and to compare the total information obtained by the two methods, the trajectory map presented in Fig. 4 was obtained. In general, for the three samples, it was observed that the trajectories obtained by T–I were broader than those obtained by PP. For sample CH-3 the trajectories obtained by both methods were highly comparable. This sample was described with certain hardness at the beginning, then an increase of sweetness and mint, to finish with a higher amount of freshness and smaller hardness; showing with both methods a bi-dimensional evolution. As for samples CH-1 and CH-2, there was less agreement, particularly over the first 4 time points, being the evolution shown by the PP done by consumers mostly one-dimensional. These results will be further discussed in the following section.

Table 2
ANOVA results for progressive profile parameters.

Attribute	Sample *F*-values	Attribute	Sample *F*-values
Mint I_{max}	4.30*	Sweetness I_{max}	26.21***
Mint Ti_{max}	1.83	Sweetness Ti_{max}	2.18
Mint AUC	4.46**	Sweetness AUC	25.98***
Mint R_{dec}	0.10	Sweetness R_{dec}	2.75
Mint R_{inc}	1.10	Sweetness R_{inc}	1.80
Freshness I_{max}	13.23***	Hardness I_{max}	15.96***
Freshness Ti_{max}	2.65	Hardness Ti_{max}	7.26**
Freshness AUC	18.04***	Hardness AUC	24.80***
Freshness R_{dec}	2.08	Hardness R_{dec}	10.14***
Freshness R_{inc}	3.28*	Hardness R_{inc}	0.23

*** $p < 0.001$.
** $p < 0.01$.
* $p < 0.05$.

4. Discussion

Results on chewing-gum samples description by a trained panel performing a T–I profile and a consumer's panel doing a PP have been presented allowing for the comparison not only of two different sensory tools but of two descriptive methods performed by two different types of panels in different testing conditions (sensory lab vs. *in-home*). Given the amount of variables which changed, differences (and similarities) observed must be carefully analyzed.

Concerning the analysis of the intensity vs. time curves obtained with both groups (Fig. 2a through h), the most important differences, particularly in terms of scaling, were obtained for freshness and mint (Fig. 2e–h respectively). Even if freshness is a trigeminal sensation and mint is an aroma, they are both highly related from a cognitive and a sensory point of view (Saint-Eve et al., 2010; Zellner & Durlach, 2002). Labbe, Gilbert, Antille, and Martin (2009) stated that the repeated exposure to mint aroma and cold stimulus paired during consumption may lead to a cognitive association between the combined mint/cold in mouth experience and the refreshing sensation. In addition, previous work shows that consumer perception under certain cases would typically be synthetic rather than analytical. Bingham, Birch, de Graaf, Behan, and Perring (1990) found that a consumer panel demonstrated enhancement of sucrose sweetness by maltol while for the same task no enhancement was found by a trained panel, supporting the hypothesis that an analytical approach is responsible for eliminating the enhancement effect. Moreover, not having received training and separated references (as the trained panel did for mint and fresh/cool sensation), consumers might not have been able to distinguish both sensations integrating them as a whole, particularly in terms of intensity. Nonetheless, in spite of the values given, sample ranking was quite similar for both groups. CH-3 had more freshness than the other two samples, especially

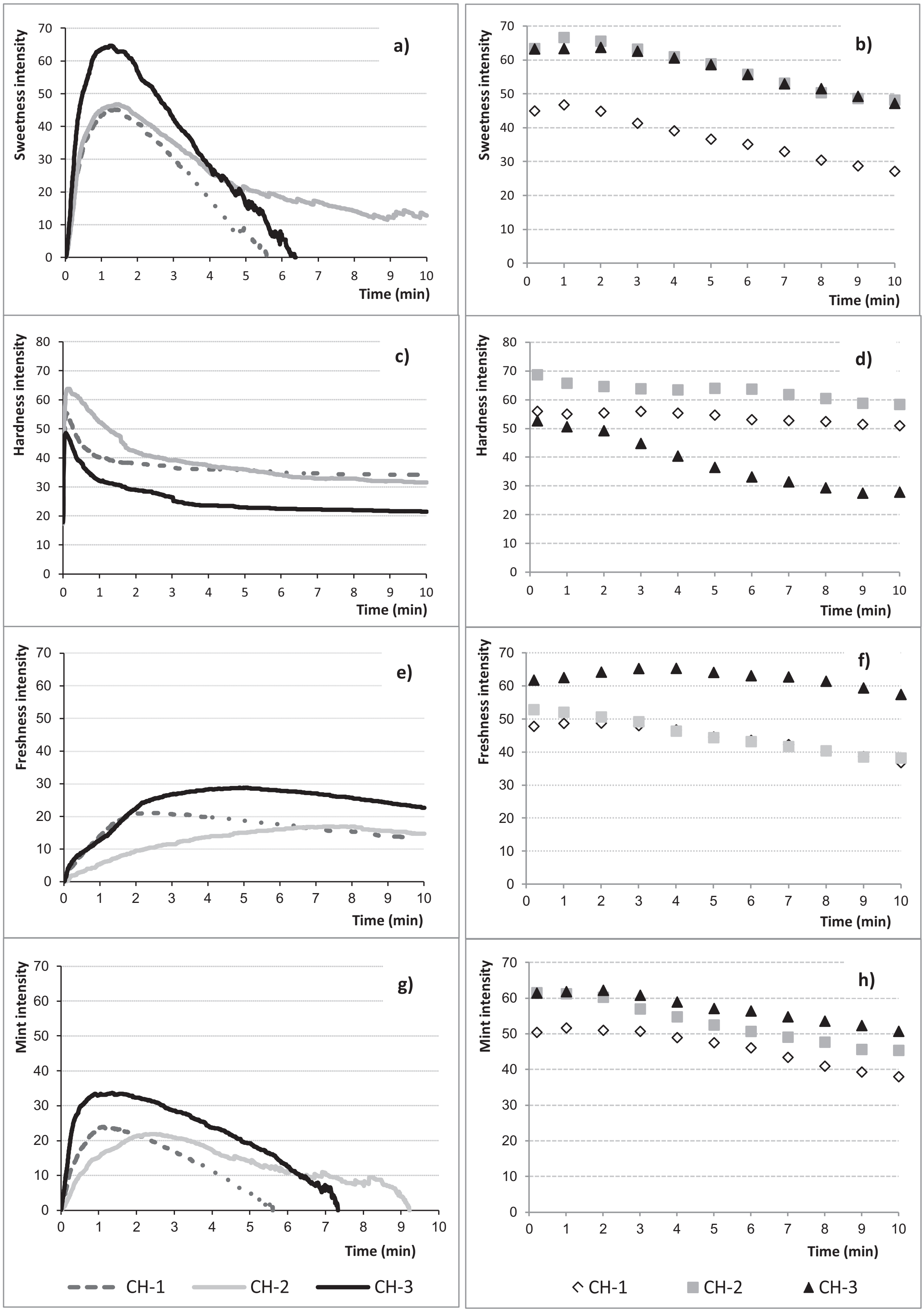

Fig. 2. Dynamic description of each chewing-gum for the attributes: sweetness (a, T–I curves; b, PP), hardness (c, T–I curves; d, PP), freshness (e, T–I curves; f, PP) and mint (g, T–I curves; h, PP).

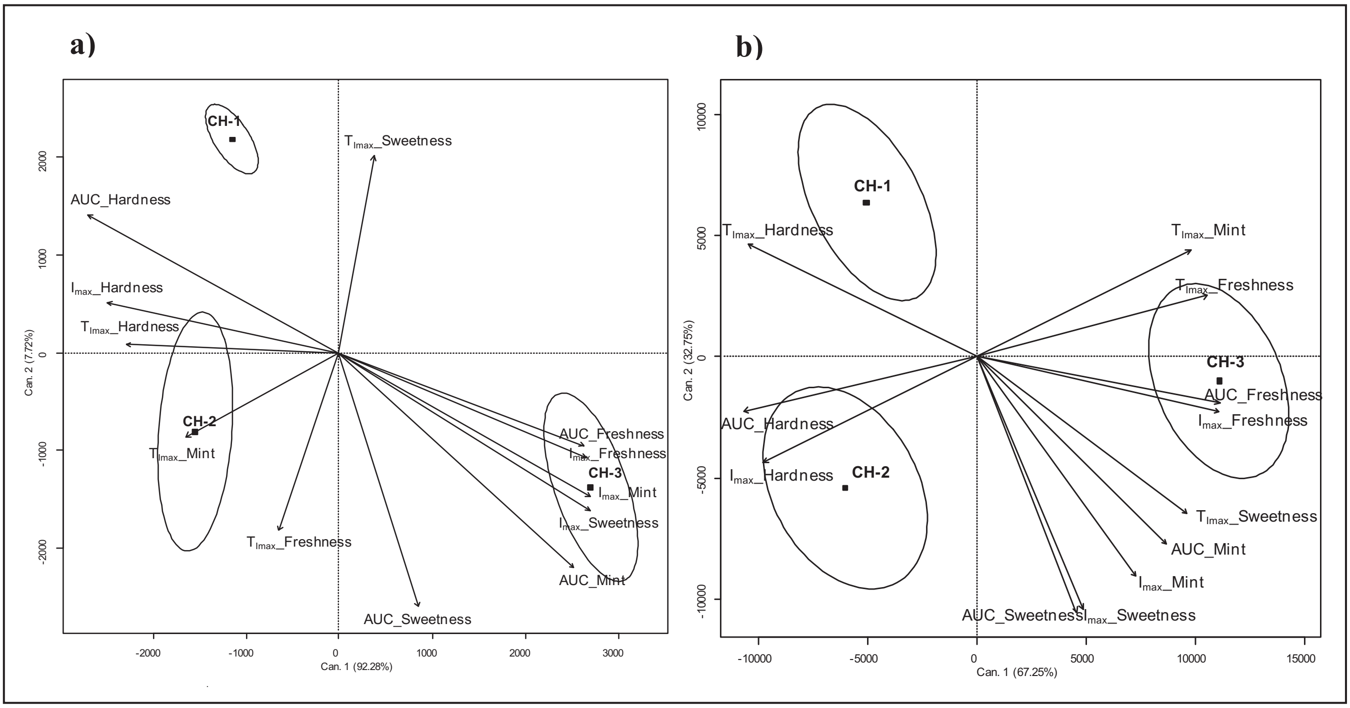

Fig. 3. (a) CVA Biplot of Time–Intensity parameters. Hotelling–Lawley MANOVA test showed significant product differences ($F = 16.934$, p-value <0.0001, NDIM = 2). 90% confidence ellipses of product means. (b) CVA Biplot of progressive profile parameters. Hotelling–Lawley MANOVA test showed significant product differences (Hotelling–Lawley MANOVA test: $F = 7.276$, p-value <0.0001, NDIM = 2). 90% confidence ellipses of product means.

after the first 5 min. In addition, sample CH-1 was the one with the least mint flavor overall.

It should also be pointed out that in terms of duration of the sensations, there were differences between the two panels for sweet and mint particularly for samples CH-1 and CH-3. While for the trained subjects the perception of these sensations declined relatively rapidly reaching zero, for the consumers the decline was moderate and did not reach zero. This could be related to the fact that in the PP consumers were able to see their previous ratings from one scoring moment to another. This was proposed for them to have an anchor while using the scale. Nonetheless, it was maybe not helpful for the repeated assessment of the sensation. For example, in order to encourage consumers to re-evaluate their perception over time during consumption, Thomas, Visalli, Cordelle, and Schlich (2015) erased from the screen their ratings after 3 s. However, in this work, authors were interested on appreciation and the number of ratings given by consumers was not fixed. In the present case, showing the rating helped not only with the use of scales but also to avoid missing data.

Some of these mentioned differences were also evidenced on the trajectory map (Fig. 4) where trajectories obtained by T–I were broader than those obtained by PP, particularly for two of the samples (CH-1 and CH-2). Thus reflecting the precision and the highest amount of information that can be obtained by T-–I. This is probably related to the highest reactivity of a continuous dynamic method – particularly at the beginning of the tasting (time points 1–3) – as opposed to a discontinuous one such as PP. Samples CH-1 and CH-2 did not evidence much change in hardness along time according to the PP while by T–I some changes were observed between time zero and 2 min. This is probably why trajectories for PP were mostly one-dimensional, showing mostly changes in sweetness and mint. Since changes in hardness were more evident for sample CH-3, resulting trajectories were more comparable for this sample.

When comparing the parametric results obtained by both panels (Fig. 3a and b) it could be observed that both panels were in agreement for the AUC and I_{max} of the four attributes. This agreement for the AUC parameter is showing that in a global manner, the temporal impression of each chewing-gum considering the whole consumption period was similar for both groups and by the two methodologies. In this way, sample CH-3 was strongly characterized by the I_{max} and AUC of freshness, mint and sweetness according to the two different groups. In samples CH-1 and CH-2 hardness was an important descriptor (also indicated by parameters AUC and I_{max}).

The most important disagreement between the two groups was in relation to T_{Imax} in all attributes. This is more likely related to the method (T–I vs. PP) than to the fact that the task was developed by consumers or by a trained panel. According to Devezeaux de la vergne et al. (2015), PP can be used instead of T–I when information about dynamic changes in perception are investigated without the need of a high time resolution. Therefore, during the first part of the evaluation, where changes might happen at a higher speed rate, differences between both methods could be accentuated. The most similar information for T_{Imax} between both methods was obtained for the attribute hardness. The time it takes for a chewing-gum to reach its maximum intensity of hardness is big enough (min instead of sec or even millisec) to make it possible to obtain roughly the same values with a continuous (T–I) and discrete (PP) method. Nonetheless, it should be pointed out that in the case of these products, T_{Imax} was altogether not a highly differentiating parameter.

So T–I would be giving somewhat richer information, but the general description done by consumers using PP also gives coherent descriptive data. As for T–I, one of the discussed issues of the T–I method is the dumping effect which occurs when a single attribute within a food is measured since this would be rated as more intense when evaluated alone than when evaluated with addi-

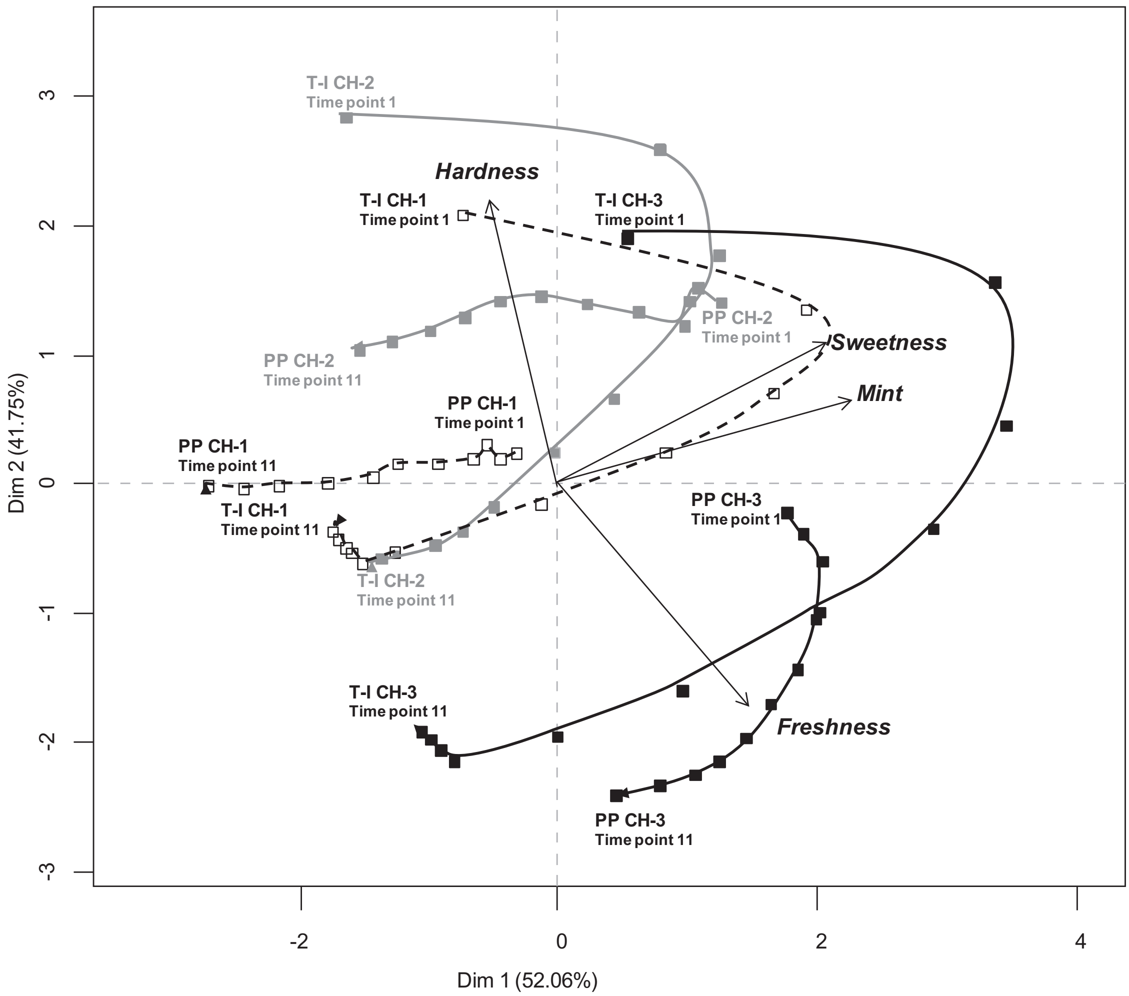

Fig. 4. Principal Component Analysis. Biplot representing the sensory trajectories of the three samples evaluated by both methods (Time–Intensity and progressive profile).

tional attributes (Devezeaux de la vergne et al., 2015; Duizer et al., 1997). This would not be the case of PP were the option of quantifying every attribute is available at all times. However, Clark and Lawless (1994) noted that prior knowledge of which attributes are to be rated should abolish any halo-dumping effects even if these are rated in succession since the subject realizes that there is no need to "dump" qualities (Prescott, 1999). Also, according to van der Klaauw and Frank (1996) providing appropriate response alternatives encourages observers to separate the component attributes of a complex stimulus, whereas they are most likely to integrate dimensions when their response alternatives are limited. In the case of the present work, the trained panel was aware of the four different attributes they had to measure along sessions avoiding this potential disadvantage of the T–I method.

Taking all this into account, when making a methodological choice it is also important to evaluate the ratio of time invested vs information obtained. In terms of costs and time consumption, the PP here done by consumers represented 30 min net of work for each consumer, resulting in a total of 25 h of human work which, in addition, could be done in parallel thanks to data acquisition in *in-home* conditions. On the other hand, doing the T–I profile with the trained panel represented 9 panelists who worked 8 h net each (materials and methods section) plus the hours of training; meaning a total of approximately 144 h of human work, nearly 6 times as much than the PP in *in-home* conditions with consumers. In addition, since these working hours are distributed only among 9 people, it meant that the experiment took several weeks to be completed as opposed to the 3 days needed with the other method. Needless to say, this efficiency and time consumption can be directly translated into costs. So even though the information obtained by the PP done by consumers might be less rich, the reduction of the time needed for the evaluation of samples approximately by a factor of ten provides a new tool which could be highly appreciated when working, for example, in product development.

5. Conclusion

From a general perspective, both methods gave similar information. The global impression of each chewing-gum was well described by consumers in relation to the results obtained by the trained panel. By using a PP some information was lost, but this was probably linked to the method itself and not to the fact that it was done by consumers. From a methodological point of view, the same PP could also have been done with the trained panel in order to compare results. However, given the lack of precision in the method itself, doing it with a large number of consumers can increase its reliability.

According to present results, T–I method might still be recommended in order to evaluate small changes in formulation. Even if more laborious, this technique gives richer information and might be able to point out small differences. However, if looking to validate the sensory description of a new different prototype, the PP done by consumers in *in-home* conditions might be a highly interesting option being cost and time efficient.

References

Bingham, A. F., Birch, G. G., de Graaf, C., Behan, J. M., & Perring, K. D. (1990). Sensory studies with sucrose–maltol mixtures. *Chemical Senses, 15*, 447–456.

Clark, C. C., & Lawless, H. T. (1994). Limiting response alternatives in time–intensity scaling: An examination of the halo-dumping effect. *Chemical Senses, 19*, 583–594.

Cliff, M., & Heymann, H. (1993). Development and use of time–intensity methodology for sensory evaluation: A review. *Food Research International, 26*, 375–385.

Devezeaux de la vergne, M., Van delft, M., Van develde, F., Van boekel, M. A. J. S., & Stieger, M. (2015). Dynamic texture perception and oral processing of semi-solid food gels: Part 1: Comparison between QDA, progressive profiling and TDS. *Food Hydrocolloids, 43*, 207–217.

Duizer, L. M., Bloom, K., & Findlay, C. J. (1997). Dual-attribute time–intensity sensory evaluation: A new method for temporal measurement of sensory perceptions. *Food Quality and Preference, 8*, 261–269.

Eilers, P. H. C., & Dijksterhuis, G. B. (2004). A parametric model for time–intensity curves. *Food Quality and Preference, 15*, 239–245.

Galmarini, M. V., Symoneaux, R., Visalli, M., Zamora, M. C., & Schlich, P. (2015). Static vs. dynamic liking in chewing gum: A new approach using a background task and a natural setting. *Food Quality and Preference, 40B*, 381–386.

Garrido, D., Calviño, A., & Hough, G. (2001). A parametric model to average time–intensity taste data. *Food Quality and Preference, 12*, 1–8.

Husson, F., Josse, J., Le, S., & Mazet, J. (2013). FactoMineR: Multivariate exploratory data analysis and data mining with R. *R Package Version, 1*, 102–123.

Husson, F., le Dien, S., & Pagès, J. (2001). Which value can be granted to sensory profiles given by consumers? Methodology and results. *Food Quality and Preference, 12*, 291–296.

ISO 1999. Guide for the characterization of the evolution of the intensity of a sensory attribute during the time. Sensory analysis. Geneva, Switzerland.

Jack, F. R., Piggott, J. R., & Paterson, A. (1994). Analysis of textural changes in hard cheese during mastication by progressive profiling. *Journal of Food Science, 59*, 539–543.

Kuesten, C., Bi, J., & Feng, Y. (2013). Exploring taffy product consumption experiences using a multi-attribute time-intensity (MATI) method. *Food Quality and Preference*, 260–273.

Labbe, D., Gilbert, F., Antille, N., & Martin, N. (2009). Sensory determinants of refreshing. *Food Quality and Preference, 20*, 100–109.

Labbe, D., Schlich, P., Pineau, N., Gilbert, F., & Martin, N. (2009). Temporal dominance of sensations and sensory profiling: A comparative study. *Food Quality and Preference, 20*, 216–221.

Lawless, H. T., & Heymann, H. (1998). *Sensory evaluation of food: Principles and practices*. International Thomson Publishing Services Ltd.

Lawless, H., & Heymann, H. (1999). Descriptive analysis. In *Sensory evaluation of food*. Springer, US.

Lee, W., III, & Pangborn, M. (1986). Time–intensity: The temporal aspects of sensory perception. *Food Technology, 40*, 78–82.

Lenzi, S., Kar, S., Michaelidou, T. A., Harvey, J. E., Beam, M. A., McCormick, D. T., ... Campomanes-Marin, J. P. (2012). Chewing gum compositions providing flavor release profiles. *EP Patent, 2*(403), 349.

Martin, C., Visalli, M., Lange, C., Schlich, P., & Issanchou, S. (2014). Creation of a food taste database using an in-home "taste" profile method. *Food Quality and Preference, 36*, 70–80.

McGowan, B. A., & Lee, S. Y. (2006). Comparison of methods to analyze time–intensity curves in a corn zein chewing gum study. *Food Quality and Preference, 17*, 296–306.

Ovejero-López, I., Bro, R., & Bredie, W. L. P. (2005). Univariate and multivariate modelling of flavour release in chewing gum using time–intensity: A comparison of data analytical methods. *Food Quality and Preference, 16*, 327–343.

Peltier, C., Visalli, M., & Schlich, P. (2015). Comparison of canonical variate analysis and principal component analysis on 422 descriptive sensory studies. *Food Quality and Preference, 40, Part B*, 326–333.

Piggott, J. R. (2000). Dynamism in flavour science and sensory methodology. *Food Research International, 33*, 191–197.

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*, 450–455.

Prescott, J. (1999). Flavour as a psychological construct: Implications for perceiving and measuring the sensory qualities of foods. *Food Quality and Preference, 10*, 349–356.

Research-and-Markets. (2014). *Chewing gums- global strategic business report*. Dublin, Ireland: Guinness Center.

Saint-Eve, A., Déléris, I., Feron, G., Ibarra, D., Guichard, E., & Souchon, I. (2010). How trigeminal, taste and aroma perceptions are affected in mint-flavored carbonated beverages. *Food Quality and Preference, 21*, 1026–1033.

Saint-Eve, A., Déléris, I., Panouillé, M., Dakowski, F., Cordelle, S., Schlich, P., et al. (2011). How texture influences aroma and taste perception over time in candies. *Chemosensory Perception, 4*, 32–41.

Song, J. H., Knutsen, H. A., Broderick, K. B., & Seielstad, D. A. (2010). Flavor releasing cores and their use in chewing gum. *EP Patent, 2*(083), 638.

Stone, H., & Sidel, J. L. (2003). *Sensory evaluation practices*. Academic Press.

Thomas, A., Visalli, M., Cordelle, S., & Schlich, P. (2015). Temporal drivers of liking. *Food Quality and Preference, 40, Part B*, 365–375.

van der Klaauw, N. J., & Frank, R. A. (1996). Scaling component intensities of complex stimuli: The influence of response alternatives. *Environmental International, 22*, 21–31.

Varela, P., & Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Research International, 48*, 893–908.

Worch, T., Lê, S., & Punter, P. (2010). How reliable are the consumers? Comparison of sensory profiles from consumers and experts. *Food Quality and Preference, 21*, 309–318.

Zellner, D., & Durlach, P. (2002). What is refreshing? An investigation of the color and other sensory attributes of refreshing foods and beverages. *Appetite, 39*, 185–186.