ASSESSMENT OF GLOBAL AND INDIVIDUAL REPRODUCIBILITY OF PROJECTIVE MAPPING WITH CONSUMERS

Leticia Vidal[1], Rafael Silva Cadena[1], Silvana Correa[2], Rosa A. Ábalos[2], Beatriz Gómez[2], Ana Giménez[1], Paula Varela[3], Gastón Ares[1]

[1] Departamento de Ciencia y Tecnología de Alimentos, Facultad de Química, Universidad de la República, Uruguay

[2] Facultad de Bromatología, Universidad Nacional de Entre Ríos, Argentina.

[3] Instituto de Agroquímica y Tecnología de Alimentos (CSIC), Avda. Agustín Escardino, 7. 46980, Paterna (Valencia), Spain

Corresponding author: Leticia Vidal

Telephone: +598 29248003

Fax: +598 292419906

Email: lvidal@fq.edu.uy

Running title: Global and individual reproducibility of projective mapping

ABSTRACT

The popularity of projective mapping with consumers for sensory characterization has markedly increased in the last five years. To have confidence in this methodology it is necessary to ensure that a similar product profile would emerge if the test was repeated. Also, deciding whether the study should be replicated or not is a key issue in test implementation. In this context, the aim of the present work was to evaluate global and individual reproducibility of projective mapping for sensory characterization with consumers and to evaluate the influence of the size of difference among samples. Six consumer studies were conducted using a test–retest paradigm. In each study, responses from the same group of consumers to the same sample set in two different sessions were compared. Across the six studies individual reproducibility tended to be low. However, the RV coefficients of consensus sample configurations between sessions were higher than 0.75, suggesting that test-retest reproducibility of projective mapping with consumers proved to be relatively high.

PRACTICAL APPLICATIONS

The present work provides evidence of the reproducibility of projective mapping for sensory characterization with consumers. Although sample configurations were stable, some differences in conclusions regarding similarities and differences among samples were identified between sessions. This indicates that care must be taken when relying on results of projective mapping with consumers obtained without the use of replicates, particularly when working with sample sets with small differences. Results from the present work showed that stability indices of sample configurations based on boostrapping resampling approaches were related to global reproducibility. These indices could be useful to decide whether or not it is necessary to replicate projective mapping in order to ensure that conclusions regarding similarities and differences among samples would be repeatedly identified. This is of particular interest considering the difficulty of asking consumers to attend separate sessions.

Keywords: sensory characterization; napping; projective mapping, consumer(s)

INTRODUCTION

Sensory characterization is one of the most powerful and extensively used tools in sensory science (Lawless and Heymann 2010). Descriptive analysis with highly trained assessors has been the most popular method for sensory characterization in the last decades (Meilgaard et al. 1999; Murray et al. 2001; Stone et al. 1974). Although this methodology provides detailed, consistent, reproducible and stable in time results, it is time consuming and can be quite expensive and difficult to apply in many situations (Murray et al. 2010; Varela and Ares 2012). Therefore, the development of simpler and faster methods which use consumers to describe products are becoming more accepted within the sensory science community and are increasingly considered a valid alternative to obtain the sensory profile of a set of products (Valentin et al. 2012; Varela and Ares 2012).

Projective mapping or Napping® is one of the novel methodologies for sensory characterization which has been increasingly used in the last five years (Varela and Ares 2012). It is a projective type method which collects bi-dimensional perceptual maps for each assessor in a single sensory session (Risvik et al. 1994). Samples are simultaneously presented, and have to be positioned by each assessor on a bi-dimensional space according to the global differences and similarities among them, in such a way that the more similar they are, the closer they should be on the provided space (Risvik et al. 1994; Risvik et al. 1997).

Projective mapping has been reported to be a simple methodology, which can be performed by trained assessors or consumers (Valentin et al. 2012; Varela and Ares 2012). It has been applied to a wide range of food products such as chocolate (Risvik et al. 1994), ewe milk cheeses (Barcenas et al. 2004), wine (Pagès 2005; Perrin et al. 2008; Ross et al. 2012), apples (Nestrud and Lawless 2010d), milk desserts (Ares et al. 2010a), fish nuggets (Albert et al. 2011) and powdered drinks (Ares et al. 2011).

It is necessary to ensure that both valid and reproducible information is provided by projective mapping before it can be established as a standard methodology for sensory characterization with consumers. If validity is taken to mean that projective mapping provides sensory characterizations similar to those from Descriptive analysis with trained assessors, then it has been already established by several authors (Louw et al. 2013, Risvik et al. 1997; Pagès 2005; Perrin et al. 2008).

Reproducibility of projective mapping has been less explored in the literature and one of the questions that arises when implementing projective mapping for sensory characterization is whether the task should be replicated or not (Hopfer and Heymann 2013). Projective mapping can be regarded as a reproducible methodology if it provides similar results when executed under identical conditions in different sessions separated in time (Yu 2005). In the great majority of studies using projective mapping assessors complete the task only once (Albert et al. 2011; Ares et al. 2010; Ares et al. 2011; Dehlholm et al. 2012a; Kennedy and Heymann 2009; King et al. 1998; Nestrud and Lawless 2008; 2010; Pagès 2005; Pagès et al. 2010; Perrin et al. 2008). In some studies the reproducibility of projective mapping has been evaluated using a blind duplicate sample within the same session (Moussaoiu and Varela 2010; Nestrud and Lawless 2008; 2010; Veinand et al. 2011). Only few studies have reported repeated evaluations of projective mapping (Barcenas et al. 2004; Hopfer and Heymann 2013; Kennedy 2010; Perrin and Pagès 2009; Risvik et al. 1994; 1997). At the individual level, Kennedy (2010) and Risvik et al. (1994; 1997) have reported low reproducibility which have been attributed to changes in consumer arrangement criteria. In particular, Kennedy (2010) reported that most consumers showed an RV coefficient lower than 0.5 for three replicated sample configurations of granola bars. However, at the aggregate level most studies have shown that consensus sample configurations and conclusions regarding overall similarities and dissimilarities among the samples are very similar across replicates (Hopfer and Heymann 2013; Kennedy 2010; Perrin and Pagès 2009; Risvik et al. 1994; 1997). Barcenas et al. (2004) reported some changes in sample configurations from triplicate evaluations of ewes milk cheeses. However, the authors could not explain if these differences were due to changes in assessors' perception or to changes in processing conditions which modified the sensory characteristics of the samples.

Considering that in many situations it is not practical to recruit consumers for replicate evaluations, the reproducibility of projective mapping in consumer studies deserves further exploration to ensure that reliable information can be gathered without the use of replicates. In this context, the aim of the present work was to evaluate global and individual reproducibility of projective mapping with consumers and to assess how they would be affected by the degree of differences among samples.

MATERIALS AND METHODS

Six consumer studies were conducted using a test–retest paradigm to assess individual and global reproducibility of projective mapping. In each study, responses from the same group of respondents to the same sample set in two different sessions were compared. Studies 1 and 2 required consumers to evaluate crackers in two sessions separated 48 hours, while in Studies 3-6 consumers evaluated vanilla milk desserts in two sessions held 2 weeks apart. In both cases the time between replicates was enough to assure that participants would not remember their responses from the previous session. Different times between replicates were considered to provide greater robustness to the findings.

Studies 1 and 2

Samples

Sixteen commercial brands of plain crackers (named A–P), available in the Argentinean market were evaluated. Two sets of 8 plain crackers were considered with varying degree of difference among samples: one set with large differences among 4 salted - I to L - and 4 unsalted - M to P - crackers (Study 1), and a second one with smaller differences among samples, using salted plain crackers only - A to H - (Study 2).

Participants

One hundred and eighty participants were recruited among students and workers of the Facultad de Bromatología of Universidad Nacional de Entre Ríos (Gualeguaychú, Argentina). Their ages ranged from 16 to 63 years and 73% were female. Consumers were randomly divided into 2 groups: 89 consumers participated in Study 1, while 91 consumers participated in Study 2. Consumers evaluated the sample set of each study in two separate sessions, 48 hours apart. They signed an informed consent agreement.

Data collection

For each study, consumers evaluated eight samples using a projective mapping task followed by a description phase in each session. Consumers were asked to try the eight samples and to place them on an A3 white sheet (42 x 30 cm), according to their similarities or dissimilarities (similar samples should be located close, while different samples should be located far from each other). They were explained that they had to complete the task according to their own criteria and that there was no right or wrong answers. After positioning the samples consumers were asked to provide a description of the samples. Testing took place in a sensory laboratory in individual sensory booths, designed in accordance with ISO 8589 (1988). Artificial daylight, constant temperature (22ºC) and air circulation were controlled. Still mineral water was available for rinsing.

Studies 3 to 6

Samples

Eight samples of vanilla milk desserts were prepared for each study varying in degree and type of differences among samples. Samples in Study 3 (named A1 - A8) and Study 5 (named C1 - C8) only differed in flavor, while samples of Study 4 (named B1 – B8) and 6 (named D1 – D8) presented both flavor and texture differences. Additionally, based on sample formulations, Studies 3 and 4 involved the evaluation of samples with large differences among them, while in Studies 5 and 6 differences among samples can be regarded as small. The formulation of the milk desserts is shown in Table 1 of the supplementary material section.

Desserts were prepared by mixing the solid ingredients with water and poured into a Thermomix TM 31 (Vorwerk Mexico S. de R.L. de C.V., México D.F., México). The dispersion was heated at 90ºC for 5 min under strong agitation (1100 rpm). The desserts were placed in closed glass containers, cooled to room temperature (25ºC) and then stored refrigerated (4–5ºC) for 24 h prior to their evaluation.

Participants

Four different groups of consumers were recruited among students and workers of the Facultad de Quimica of the Universidad de la República (Montevideo, Uruguay). Participants ranged in age from 20 to 50 years old and approximately 60% were female. Two groups of 48 consumers participated in Studies 3 and 4, while Studies 5 and 6 were carried out with two groups of 42 consumers. In each study, consumers participated in two separate sessions, 14 days apart. They signed an informed consent agreement and were given a small present for their participation.

Data collection

For each of the four studies (Studies 3-6), consumers evaluated eight samples of each set using a projective mapping task followed by a description phase in each session. Consumers received 15g of each vanilla milk dessert coded with 3-digit random numbers at 10ºC in plastic containers and a spoon. Mineral still water was available for rinsing between samples. Participants were asked to try the samples and to place them on an A3 white sheet (42 x 30 cm), according to their similarities or dissimilarities. Testing took place in a sensory laboratory in standard sensory booths that was designed in accordance with ISO 8589 (1988), under artificial daylight, temperature control (22ºC) and air circulation was controlled.

Data analysis

For each consumer map, the X and Y coordinates of each sample were determined, considering the left bottom corner of the sheet as the origin of coordinates. The X and Y coordinates for each session and sample set were analysed using Multiple Factor Analysis (MFA) (Pagès 2005). Confidence ellipses were constructed as suggested by Dehlholm, et al. (2012b).The stability of sample configurations from each session was evaluated using a bootstrapping resampling approach. According to Blancher et al. (2012), sample configurations can be regarded as stable if simulated repeated experiments provide similar results than those obtained with the original dataset. In the present work, the bootstrapping process consisted of obtaining 1000 subsets of size equal to the total number of consumers using random sampling with replacement. For each subset sample configurations were obtained using MFA and agreement between each of these configurations and the reference configuration (obtained with all the consumers who participated in the study) was evaluated by computing the RV coefficient (Abdi 2010). Average values and standard deviations over the RV coefficients were calculated. The RV coefficient has been used as a tool to assess the global similarity between two factorial configurations of the same products (Faye et al. 2004; de Saldamando et al. 2013). This coefficient takes the value of 0 if the configurations are uncorrelated and the value of 1 if the configurations are homothetic. It depends on the relative position of the points in the configuration and therefore is independent of rotation and translation (Robert and Escoufier 1976).

The similarities among the sample configurations over all assessors and sessions were evaluated with the RV coefficient. Also, RV coefficients of individual sample configurations between sessions were calculated as a measure of individual reproducibility. The significance of the RV coefficient was tested using a permutation test, as suggested by Josse et al. (2008). If the RV coefficient between two sample configurations is significant, it can be concluded that they are correlated and therefore information about the similarities and differences among samples is similar.

The words elicited by consumers in the description phase were qualitatively analyzed. Words with similar meaning were grouped into categories and their frequency of mention was determined by counting the number of consumers who elicited words within each category. Terms mentioned by at least 5% of the consumers were retained for further analysis (Symoneaux et al. 2012). In each session, consensual terms were identified using the methodology proposed by Kostov et al. (2013). Consensual terms were identified as those for which the p-value, computed as the proportion of random subsets, selected following a bootstrap methodology, having a within-inertia smaller or equal to the observed inertia, was smaller than 0.10. Multiple factor analysis for contingency tables (MFACT) was applied on the frequency table of each session to obtain a representation of terms (Bécue-Bertaut and Pagès 2004). In this analysis only the terms used by consumers in both sessions were considered.

All statistical analyses were performed with R language (R Development Core Team 2007) using FactoMineR (Lê et al.2008) and SensoMineR (Lê and Husson 2008).

RESULTS

Global reproducibility

No differences were observed in the percentage of inertia explained by the first and second dimensions of the MFA between sessions (Figures 1 and 2). Average RV coefficient across simulations from the bootstrapping resampling approach did not vary between sessions, suggesting that duplicate evaluation did not increase the stability of sample configurations (Table 1). As expected, average RV coefficient increased with the size of difference among samples, i.e. it was higher for the studies with large differences among samples than for studies with small differences among samples. Besides, the stability of sample configurations for the studies which included samples with flavor and texture differences was higher than that of the studies which only included flavor differences (Table 1).

- Please insert Table 1 around here-

At the aggregate level the RV coefficient of sample configurations from different sessions was higher than 0.75 (Table 2), providing evidence for the global reproducibility of projective mapping. As expected, global reproducibility increased with the size of differences among samples, as denoted by the increase in RV coefficient of sample configurations between sessions. Besides, when small differences among samples were considered, consumers were more reproducible when evaluating samples with texture and flavor differences. As shown in Table 2, the RV coefficient of sample configurations was higher for Study 6 than for Study 5.

- Please insert Table 2 around here-

Despite the high similarity in sample configurations between sessions, some differences in conclusions regarding similarities and differences among samples were identified in some of the studies. Although the RV coefficient of sample configurations between sessions for Study 1 was 0.96, the position of sample I clearly differed (Figure 1(a)). In the first session, sample I was located in a distinct position in the first and second dimensions of the MFA, whereas in the second session it was regarded as largely similar to samples L and J (their confidence ellipses overlapped). A similar difference was observed in the position of sample H in Study 2 (Figure 1(b)). Studies 4 and 6 showed highly similar sample configurations in both sessions (Figures 2(b) and 2(d)), with no differences in relation to the confidence ellipses that overlapped. The fact that samples differed in texture could have helped consumers to locate samples more easily and more reproducibly. In Studies 3 and 5 several differences can be identified in the relative positioning of the samples and consequently in the conclusions regarding similarities and differences among samples (Figures 2(a) and 2(c)).

- Please insert Figure 1 around here-

- Please insert Figure 2 around here –

Consumer descriptions

As shown in Table 3, for the six studies the number of terms used for describing samples in the description phase of projective mapping was similar for session 1 and 2 and the majority of the terms were used in both sessions. This provides preliminary evidence of the stability of consumer descriptions. The terms used in both sessions of the six studies for describing samples are shown in Table 2 of the supplementary material section.

For each study, consensual terms for a significance level of $p \leq 0.10$ were determined following the methodology proposed by Kostov et al. (2013). For all the studies the number of consensual terms was markedly lower than the total number of terms used for describing samples (Table 3). It is interesting to note that for Studies 3-6, the number of consensual terms was higher for the second session than for the first session. Besides, the number of consensual words tended to increase with the size of difference among samples.

The majority of the consensual terms identified in the first session were also consensual in the second session. For example, 6 of the 8 consensual terms identified in the first session of Study 3 were also consensual in the second session (*Caramel flavour, Consistent, Not much flavour intensity, Not very sweet, Vanilla flavour,* and *Very sweet*) (Table 2 of the supplementary material section). On the other hand, none of the consensual terms identified in the first session of Study 5 were consensual in the second session, which could be related to the fact that samples had small flavor differences.

- Please insert Table 3 around here-

MFACT allows the visualization of the descriptors used by consumers to describe samples in the two sessions of the 6 studies (Figure 3). Identical terms are connected with a line to indicate the size of the difference in how the term was used between the sessions. The terms used for describing samples differed in their reproducibility. Some of the terms were used in a markedly similar way in both sessions, being located close to each other in the first and second session. In general, the most stable terms were those which described the main sensory differences among samples. For example, in Study 1 the terms *Salty, No salt, Toasted, Burnt, Not toasted, and Crunchy* were highly reproducible (Figure 3(a)). Something similar was observed in Study 6 with the terms *Liquid, Runny, Consistent, Thick, Viscous, Creamy, Sweet* and *Very sweet* (Figure 3(f)).

On the other hand, terms describing complex sensory properties or characteristics of the desserts that did not vary among samples tended to be less stable. For example, in Study 3, which included samples with flavor differences but with the same texture, the terms *Consistent, Creamy* and *Smooth* were unstable, together with complex flavor attributes as *Aftertaste*, *Cookie* and *Milky flavor* (Figure 3(c)) The rest of the terms, particularly those related to flavor differences (e.g. *Caramel flavor, Vanilla flavor, Very Sweet, Sweet, Not sweet,* and *Not very sweet*), were located close to each other, suggesting high reproducibility in how consumers described samples across sessions. Similarly, the least reproducible terms in Study 5 were mainly related to texture characteristics which did not differ across samples (*Smooth, Thick*) and complex flavor terms (*Artificial flavor, Tasty*) (Figure 3(e)). The reproducibility of the terms depended on the size of difference among samples. Consumers tended to be more reproducible when describing samples with large differences (Figure 3(a), 3(c) and 3(d)) than when describing sample sets with small differences (Figure 3(b), 3(e) and 3(f)). Besides, in the milk dessert experiments (studies 3-6) consumers were more reproducible in describing samples with texture and flavor differences than samples that only differed in their flavor characteristics (c.f. Figures 3(c), (d), (e) and (f)).

- Please insert Figure 3 around here-

The terms that were consensual in both sessions tended to be highly reproducible between sessions (Figure 3), suggesting that the terms that were used similarly by consumers were also used in the same way over sessions. However, it is interesting to note that the most reproducible terms were not necessarily consensual in both sessions. Many terms that were used in a highly reliable way in both sessions were not consensual in any of the sessions. For example, as shown in Figure 3(a) the term *No salt* was reliably used in Study 1 but was not consensual in any of the sessions. On the contrary, the terms *Toasted flavor* and *Bitter* were among the least reproducible while they were consensual in one of the sessions.

The RV coefficients between the frequency tables of both sessions tended to be high, reaching values higher than 0.80 (Table 3). These results suggest that although some of the terms were not reliably used between sessions, descriptions obtained in both sessions provided similar information regarding similarities and differences among samples. As expected, RV coefficient between the frequency tables of consumer descriptions increased with the size of differences among samples, reaching values higher than 0.94 for the studies which included large differences among samples (Table 3).

Consumer individual reproducibility

Although global reproducibility was high, consumer individual reproducibility tended to be low (Table 2). The RV coefficients of individual sample configurations between sessions ranged from 0.001 to 0.975, indicating large differences among consumers' performance. However, average consumer reproducibility was low, as well as the percentage of consumers whose configurations were significantly correlated. For 4 out of the 6 studies less than 50% of the consumers sample configurations were significantly correlated.

As expected, consumer individual reproducibility markedly increased with the size of the differences among samples. For example, average RV coefficient of individual configurations was 0.52 for milk dessert samples with large flavor differences (Study 3) and 0.26 for samples with small flavor differences (Study 5). Additionally, in these studies the percentages of consumers whose configurations were significantly correlated between sessions were 54% and 18%, respectively (Table 2).

DISCUSSION

The present work evaluated global and individual reproducibility of projective mapping for sensory characterization with consumers using samples sets that differed in the size of difference among samples. Across the six studies, the RV coefficients of sample configurations between sessions were higher than 0.75. The minimum RV value that has been considered as indicator of good agreement between sample configurations ranges from 0.65 to 0.85 (Abdi et al. 2007; Faye et al. 2004; Kennedy 2010; Lawless and Glatter 1990; Lelièvre et al. 2008). Considering these values it can be concluded that in the present study sample configurations were relatively stable across sessions and that in the six studies test-retest reproducibility of projective mapping with consumers proved to be relatively high. These results are in agreement with several authors that reported that consensus sample configurations from projective mapping with trained and untrained assessors were stable across sessions (Hopfer and Heymann 2013; Kennedy 2010; Perrin and Pagès 2009; Risvik et al. 1994; 1997). High reproducibility of consumer-based sensory characterization has also been reported for other methodologies like sorting tasks (Cartier et al. 2006; Chollet et al. 2011; Lawless and Glatter 1990) and check-all-that-apply questions (Jaeger et al. 2013).

Despite the fact that RV coefficients were higher than 0.75, some differences in conclusions regarding similarities and differences among samples were identified between replicates, particularly for studies which involved samples with small differences. A similar result has been reported by Barcenas et al. (2004) when working with ewes' milk cheeses. These authors reported that the relative position of two samples changed across replicates, modifying conclusions regarding their similarities and differences with the rest of the sample set. On the contrary, Kennedy (2010) and Hopfer and Heymann (2013) reported that overall similarities and dissimilarities among the samples were stable over the triplicate evaluation. Results from the present work suggest that for sample sets with small differences care must be taken when drawing conclusions from sample configurations obtained using projective mapping with consumers without the use of replicates. Further research is necessary to determine if replicated projective mapping is necessary prior to the design of the study.

In the present work the majority of the terms elicited to describe samples in the description phase of projective mapping were used in a similar way in both sessions (Figure 3). Overall, the terms responsible for the main differences in the sensory characteristics of the samples were highly reproducible, while terms related to complex sensory attributes or characteristics that did not differ among samples tended to be not reproducible. This suggests that consumer descriptions in projective mapping tasks should be taken with care, particularly when evaluating samples with small differences. Although open-ended questions have been considered as an alternative method for sensory characterization with consumers (Ares et al., 2010b; Symoneaux et al., 2012), results from the present work show that consumers are not reproducible when using many terms. This would suggest the need to check the reliability of the terms for concluding on the main sensory characteristics responsible for similarities and differences among samples.

Methodologies which enable the selection of reliable terms would be useful to improve the interpretation of sensory spaces obtained from the application of holistic methodologies with consumers. Kostov et al (2013) proposed the identification of consensual terms for selecting the most reliable terms elicited in free description tasks. In the present work this methodology was not able to predict the reproducibility of the terms. Although consensual words in both sessions were used in a reproducible way, there were many terms that were not consensual but reproducible, as well as terms that were consensual in one of the sessions but were not reproducible. Thus, further research is needed to improve the interpretation of consumer responses to free description tasks.

Although global reproducibility was high, consumer individual reproducibility tended to be low in the six studies (Table 2). The average RV coefficients between sample configurations of the two sessions were lower than 0.55, while the percentage of consumers with significant RV coefficient between sessions was lower than 54%. This result is in agreement with Risvik et al. (1994; 1997), Barcenas et al. (2004), Hopfer and Heymann (2013) and Kennedy (2010). In particular, this last author reported that 10 out of 15 consumers had RV coefficient between replications lower than 0.5. Similar results have been reported for check-all-that-apply (CATA) questions for sensory characterization. Jaeger et al. (2013) reported that despite the fact that global reproducibility of CATA questions was high, consumer individual reproducibility tended to be low. This suggests that differences in individual performances between sessions tend to compensate among consumers, yielding stable consensus configurations.

The low RV coefficients between individual sample configurations can be attributed to differences in consumers' criteria for placing the samples, particularly due to training and familiarization with projective mapping and the sample set. In this sense, Kennedy (2010) reported that the internal consistency and agreement of untrained consumers when using projective mapping increased over triplicate evaluations. In the present work the percentage of variance explained by the first and second dimensions of the MFA and the stability of sample configurations (as evaluated through a resampling bootstrapping approach) did not increase with duplicate evaluation. However, the number of consensual terms tended to be larger in the second session than in the first one, which suggests that familiarization with the sensory space can improve consumer performance in descriptive tasks. Therefore, considering these results it would be interesting to study if familiarization with projective mapping and/or with the sample set increases assessor reproducibility when using projective mapping for sensory characterization, particularly considering that some consumers can find this methodology difficult to apply (Nestrud and Lawless 2008; Veinand et al. 2011). Several authors have included a short introduction or training prior to the projective mapping task (Barcenas et al. 2004; Carrillo et al. 2012; Hopfer and Heymann 2013; Risvik et al. 1994; 1997; Veinand et al. 2011), which can contribute to improve consumers' performance.

Global and individual reproducibility of projective mapping increased with the size of differences among samples. This observation, together with the fact that conclusions regarding similarities and differences among samples were not stable in some cases, indicates the need to define stability indices for sample configurations. These indices could be useful to decide whether or not to replicate projective mapping in order to ensure that conclusions regarding similarities and differences among samples would be repeatedly identified. Further research is necessary to determine if increasing the number of consumers can be an alternative approach to replicated evaluations for the stabilization of sample configurations. This is an interesting idea to explore considering that in many situations it is not practical to get the same consumers to repeat the study.

Studying the stability of sample configurations by sub-sampling using bootstrapping approaches could be an interesting approach and can contribute to development of guidelines for practitioners. In the present study the stability of sample configurations was studied using simulated repeated experiments by sampling repeatedly from the population of interest, as proposed by Faye et al. (2006) and Blancher et al. (2012) for sorting tasks. As shown in Tables 1 and 2, there was a good agreement between the stability and reproducibility of sample configurations. The studies which showed average RV coefficients across replications higher than 0.95 (studies 1, 3, 4 and 6) were highly reproducible, reaching RV coefficients between replicates higher than 0.90. These results suggest the need to further study the relationship between the stability and reproducibility of sample configurations from projective mapping. This type of research can contribute to the definition of threshold for deciding if results from projective mapping are reliable and whether or not replication is needed. When the stability of sample configuration is found to be low, replication of the study would be recommended to check that similarities and differences among samples remain when repeating the whole study. When replicating projective mapping tasks, conclusions should be drawn from consensus sample configurations across replicates from Hierarchical Multiple Factor Analysis (Le Dien and Pagès 2003). This methodology is an extension of MFA and balances the relevance of groups of variables with different hierarchy and provides an overall result. In the context of replicated projective mapping tasks HMFA provides consensus sample configurations after balancing data from each separate session.

CONCLUSIONS

Results from the present work showed that although most consumers were only slightly reproducible, global configurations from projective mapping were reasonably stable across sessions. Descriptions of samples were used in a similar way in both sessions, the terms responsible for the main differences were highly reproducible, while complex sensory attributes or characteristics that did not differ among samples tended to be not reproducible.

The degree (large or small) and type (flavor or flavor and texture) of difference among samples had a strong influence on both global and individual reproducibility of projective mapping, suggesting that care must be taken when relying on results of projective mapping with consumers obtained without the use of replicates. In this sense, the use of indices that evaluate the stability of sample configurations can contribute to decide whether or not a replication is needed. In the present work the stability index calculated using a boostrapping resampling approach was strongly related to consumer global reproducibility. Research in this area could contribute to the selection of criteria for evaluating the reliability of sensory characterization with consumers and to define the need of using replicates with trained, semi-trained and untrained assessors. Besides, further research on the reproducibility of projective mapping when working with samples sets of different complexity can help to decide if replicated projective mapping is necessary prior to the design of the experiment.

ACKNOWLEDGMENTS

The authors are indebted to Comisión Sectorial de Investigación Científica (Universidad de la República, Uruguay) for financial support, to Agencia Nacional de Investigación e Innovación (ANII, Uruguay) for the scholarship granded to author Leticia Vidal and to CAPES-UdelaR for the scholarship granted to author Rafael Silva Cadena.

The authors would also like to thank the Spanish Ministry of Science and Innovation for the contract awarded to the author P. Varela (Juan de la Cierva Program) and to the Spanish Ministry of Education, Culture and Sports for the José Castillejo grant awarded to author P. Varela.

REFERENCES

ABDI, H. 2010. Congruence: Congruence coefficient, RV coefficient, and Mantel Coefficient. In *Encyclopedia of Research Design* (N.J. Salkind, D.M. Dougherty and B. Frey, eds.) pp. 222-229. Sage, Thousand Oaks, CA.

ABDI, H., VALENTIN, D., CHOLLET, S. and CHREA, C. 2007. Analyzing assessors and products in sorting tasks: DISTATIS, theory and applications. Food Qual. Prefer. *18*, 627–640.

ALBERT, A., VARELA, P., SALVADOR, A., HOUGH, G. and FISZMAN, S. 2011. Overcoming the issues in the sensory description of hot served food with a complex texture. Application of QDA®, flash profiling and projective mapping using panels with different degrees of training. Food. Qual. Prefer. *22*, 463–473.

ARES, G., DELIZA, R., BARREIRO, C., GIMÉNEZ, A. and GÁMBARO, A. 2010a. Comparison of two sensory profiling techniques based on consumer perception. *Food. Qual. Prefer. 21*, 417–426.

ARES, G., GIMÉNEZ, A., BARREIRO, C. and GÁMBARO, A. 2010b. Use of an open-ended question to identify drivers of liking of milk desserts. Comparison with preference mapping techniques. Food Qual. Prefer. *21*, 286–294.

ARES, G., VARELA, P., RADO, G. and GIMENEZ, A. 2011. Are consumer profiling techniques equivalent for some product categories? The case of orange-flavored powdered drinks. International J. Food Sci. Technol. *46*, 1600–1608.

BARCENAS, P., PÉREZ ELORTONDO, F. J. and ALBISU, M. 2004. Projective mapping in sensory analysis of ewes milk cheeses: A study on consumers and trained panel performance. Food Res. Int. 37, 723–729.

BÉCUE-BERTAU, M. and PAGÈS, J. 2004. A principal axes method for comparing contingency tables: MFACT. Comput. Stat. Data Anal. *45*, 481–503.

BLANCHER, G., CLAVIER, B., EGOROFF, C., DUINEVELD, K. and PARCON, J. 2012. A method to investigate the stability of a sorting map. Food. Qual. Prefer. *23*, 36-43.

CARRILLO, E., VARELA, P. and FISZMAN, S. 2012. Packaging information as a modulator of consumers' perception of enriched and reduced-calorie biscuits in tasting and non-tasting tests. Food. Qual. Prefer. *25*, 105–115

CARTIER, R., RYTZ, A., LECOMTE, A., POBLETE, E., KRYSTLIK, J., BELIN, E. and MARTIN, N. 2006. Sorting procedure as an alternative to quantitative descriptive analysis to obtain a product sensory map. Food. Qual. Prefer. *17*, 562–571.

CHOLLET, S., LELIÈVRE, ABDI, H. and VALENTIN, D. 2011. Sort and beer: Everything you wanted to know about the sorting task but did not dare to ask. Food. Qual. Prefer. *22*, 507–520.

DE SALDAMANDO, L., DELGADO, J., HERENCIA, P., GIMÉNEZ, A. and ARES, G. 2013 Polarized sensory positioning: Do conclusions depend on the poles? Food. Qual. Prefer. *29*, 25–32

DEHLHOLM, C., BROCKHOFF, P. B., MEJNERT, L., AASLYNG, M. D. and BREDIE, W. L. P. 2012a. Rapid descriptive sensory methods – comparison of free multiple sorting, partial napping, napping, flash profiling and conventional profiling. Food. Qual. Prefer. *26*, 267–277.

DEHLHOLM, C., BROCKHOFF, P. B. and BREDIE, W. L. P. 2012b. Confidence ellipses: A variation based on parametric bootstrapping applicable on Multiple Factor Analysis results for rapid graphical evaluation. Food. Qual. Prefer. *26*, 278–280.

FAYE, P., BRÉMAUD, D., DURAND-DAUBIN, D., COURCOUX, P., GIBOREAU, A. and NICOD, A. 2004. Perceptive free sorting and verbalization tasks with naive subjects: An alternative to descriptive mappings. Food. Qual. Prefer. *15*, 781–791.

FAYE, P., BRÉMAUD, D., TEILLET, E., COURCOUX, P., GIBOREAU, A. and NICOD, H. 2006. An alternative to external preference mapping based on consumer perceptive mapping. *Food. Qual. Prefer. 17*, 604–614

HOPFER, H. and HEYMANN, H. 2013. A summary of projective mapping observations – The effect of replicates and shape, and individual performance measurements. Food. Qual. Prefer. *28*, 164–181.

ISO. 1988. *Sensory analysis: General guidance for the design of test rooms, ISO 8589.* International Organization for Standardization, Geneve.

JAEGER, S., CHHEANG, S.L., YIN, J., BAVA, C.M., GIMENEZ, A., VIDAL, L. and ARES, G. 2013. Check-all-that-apply (CATA) responses elicited by consumers: Within-assessor reproducibility and stability of sensory product characterizations. Food. Qual. Prefer. *30*, 56-67.

JOSSE, J., PAGÉS, J. and HUSSON, F. 2008. Testing the significance of the RV coefficient. Comput Stat. Data Anal. *53*, 82–91.

KENNEDY, J. 2010. Evaluation of replicated projective mapping of granola bars. J. Sensory Stud. *25*, 672–684.

KENNEDY, J. and HEYMANN, H. 2009. Projective mapping and descriptive analysis of milk and dark chocolate. J. Sensory Stud. *24*. 220–233.

KOSTOV, B., BÉCUE-BERTAUT, M. and HUSSON, F.. 2013. An original methodology for the analysis and interpretation of word-count based methods: multiple factor analysis for contingency tables complemented by consensual words. Food. Qual. Prefer. *In press*, http://dx.doi.org/10.1016/j.foodqual.2013.06.009 .

LAWLESS, H. T. and GLATTER, S. 1990. Consistency of multidimensional scaling models derived from odor sorting. J. Sensory Stud*.* *5*, 217–230.

LAWLESS, H. T. and HEYMANN, H. 2010. *Sensory evaluation of food. Principles and practices,* 2nd edition. Springer, New York.

LÊ, S. and HUSSON, F. 2008. SensoMineR: a package for sensory data analysis. J. Sensory Stud. *23*, 14–25.

LÊ, S., JOSSE, J. and HUSSON, F. 2008. FactoMineR: An R package for multivariate analysis. J. Stat. Soft. *25*, 1–18.

LE DIEN, S. and PAGÈS, J. 2003. Hierarchical multiple factor analysis: Application to the comparison of sensory profiles. Food. Qual. Prefer. *14*, 397–403.

LELIÈVRE, M., CHOLLET, S., ABDI, H. and VALENTIN, D. 2008. What is the validity of the sorting task for describing beers? A study using trained and untrained assessors. Food. Qual. Prefer. *19*, 697–703.

LOUW, L. MALHERBE, S., NAES, T., LAMBRECHTS, M., RENSBURG, P. and NIEUWOUDT, H. 2013. Validation of two Napping® techniques as rapid sensory screening tools for high alcohol products. Food Qual. Prefer. *30*, 192–201

MEILGAARD, M. C., CIVILLE, G. V. and CARR, B. T. 1999. *Sensory evaluation techniques*, 2nd edition. CRC Press, Boca Raton, FL.

MOUSSAOUI, K. A. and VARELA, P. 2010. Exploring consumer product profiling techniques and their linkage to a quantitative descriptive analysis. Food. Qual. Prefer. *21*, 1088–1099.

MURRAY, J.M., DELAHUNTY, C.M. and BAXTER, I.A. 2001. Descriptive sensory analysis: Past, present and future. Food Res. Int. *34*, 461–471.

NESTRUD, M.A. and LAWLESS, H.T. 2008. Perceptual mapping of citrus juices using projective mapping and profiling data from culinary professionals and consumers. Food. Qual. Prefer. *19*, 431–438.

NESTRUD, M.A. and LAWLESS, H.T. 2010. Perceptual mapping of apples and chesses using projective mapping and sorting. J. Sensory Stud. *25*, 309–324.

PAGÈS, J. 2005. Collection and analysis of perceived product inter-distances using multiple factor analysis: Application to the study of 10 white wines from the Loire Valley. Food. Qual. Prefer. *16*, 642–649.

PAGÈS, J., CADORET, M. and LÊ, S. 2010. The sorted Napping: A new holistic approach in sensory evaluation. J. Sensory Stud. *25*, 637–658.

PERRIN, L. and PAGÈS, J. 2009. Construction of a product space from the ultra-flash profiling method: Application to 10 red wines from the Loire valley. J. Sensory Stud. *24*, 372–395.

PERRIN, L., SYMONEAUX, R., MAÎTRE, I., ASSELIN, C., JOURJON, F. and PAGÈS, J. 2008. Comparison of three sensory methods for use with the Napping® procedure: Case of ten wines from Loire Valley. Food. Qual. Prefer. *19*, 1–11.

R DEVELOPMENT CORE TEAM. 2007. *R: A Language and Environment for Statistical Computing.* ISBN 3-900051-07-0. R Foundation for Statistical Computing, Vienna.

RISVIK, E., MCEWAN, J. A., COLWILL, J. S., ROGERS, R. and LYON, D. H. 1994. Projective mapping: A tool for sensory analysis and consumer research. Food. Qual. Prefer. *5*, 263–269.

RISVIK, E., MCEWAN, J. A. and RODBOTTEN, M. 1997. Evaluation of sensory profiling and projective mapping data. Food. Qual. Prefer. *8*, 63–71.

ROBERT, P. and ESCOUFIER, Y. 1976. A unifying tool for linear multivariate statistical methods: the RV coefficient. Applied Stat. *25*, 257–265.

ROSS, C.F., WELLER, K.M., and ALLDREDGE, J.R. 2012. Impact of Serving Temperature on Sensory Properties of Red Wine as Evaluated Using Projective Mapping by a Trained Panel. J. Sensory Stud. *27*, 463-470.

STONE, H., SIDEL, J. L., OLIVER, S., WOOLSEY, A. and SINGLETON, R. C. 1974. Sensory evaluation by quantitative descriptive analysis. *Food Technol.* 28, 24–33.

SYMONEAUX, R., GALMARINI, M. V. and MEHINAGIC, E. 2012. Comment analysis of consumer's likes and dislikes as an alternative tool to preference mapping. A case study on apples. Food. Qual. Prefer. *24*, 59–66.

VALENTIN, D., CHOLLET, S., LELIEVRE, M. and ABDI, H. 2012. Quick and dirty but still pretty good: a review of new descriptive methods in food science. Int. J.Food Sci. Technol. 47, 1563–1578.

VARELA, P. and ARES, G. 2012. Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. Food Res. Int. *48*, 893–908.

VEINAND, B., GODEFROY, C., ADAM, C. and DELARUE, J. 2011. Highlight of important product characteristics for consumers. Comparison of three sensory descriptive methods performed by consumers. Food. Qual. Prefer. *22*, 474–485.

YU, C.H. 2005. Test-retest reliability. In *Encyclopedia of social measurement,* Vol. 3, (K. Kempf-Leonard, ed.), p. 777–784. Academic Press, San Diego, CA.

FIGURE CAPTIONS

FIGURE 1. SAMPLE REPRESENTATION ON THE FIRST AND SECOND DIMENSIONS OF MULTIPLE FACTOR ANALYSIS PERFORMED ON DATA FROM THE TWO SESSIONS CONSIDERED IN: (A) STUDY 1 (SALTED -I TO L- AND UNSALTED PLAIN CRACKERS -M TO P-) AND (B) STUDY 2 (SALTED CRACKERS -A TO H-). CONFIDENCE ELLIPSES AROUND SAMPLES WERE CREATED USING PARAMETRIC BOOTSTRAPPING.

FIGURE 2. SAMPLE REPRESENTATION ON THE FIRST AND SECOND DIMENSIONS OF MULTIPLE FACTOR ANALYSIS PERFORMED ON DATA FROM THE TWO SESSIONS CONSIDERED IN: (A) STUDY 3 (LARGE FLAVOUR DIFFERENCES), (B) STUDY 4 (LARGE FLAVOUR AND TEXTURE DIFFERENCES), (C) 5 (SMALL FLAVOUR DIFFERENCES), AND (D) 6 (SMALL FLAVOUR AND TEXTURE DIFFERENCES). CONFIDENCE ELLIPSES AROUND SAMPLES WERE CREATED USING PARAMETRIC BOOTSTRAPPING.

FIGURE 3. REPRESENTATION OF THE TERMS USED BY CONSUMERS TO DESCRIBE THE SAMPLES, ON THE FIRST AND SECOND DIMENSIONS OF THE MULTIPLE FACTOR ANALYSIS FOR THE CONTINGENCY TABLES PERFORMED ON DATA FROM THE TWO SESSIONS CONSIDERED IN: (A) STUDY (PLAIN CRACKERS, LARGE DIFFERENCES), (B) 2 (PLAIN CRACKERS, SMALL DIFFERENCES), (C) 3 (MILK DESSERTS, LARGE FLAVOUR DIFFERENCES), (D) 4 (MILK DESSERTS, LARGE FLAVOUR AND TEXTURE DIFFERENCES), (E) 5 (MILK DESSERTS, SMALL FLAVOUR DIFFERENCES), AND (F) 6 (MILK DESSERTS, SMALL FLAVOUR AND TEXTURE DIFFERENCES).. TERMS USED IN THE FIRST SESSION ARE INDICATED USING GREY DIAMONDS AND ITALIC LETTERS, WHILE TERMS USED IN THE SECOND SESSION ARE INDICATED USING BLACK DIAMONDS AND REGULAR LETTERS. TERMS

HIGHLIGHTED IN BLACK WERE CONSENSUAL FOR P≤0.10 (KOSTOV ET AL. 2013). IDENTICAL TERMS ARE CONNECTED WITH A LINE TO INDICATE THE SIZE OF THE DIFFERENCE IN HOW THE TERM WAS USED BETWEEN THE SESSIONS

TABLES

TABLE 1. AVERAGE RV COEFFICIENT OF SAMPLE CONFIGURATION ACROSS SIMULATIONS OBTAINED VIA A BOOTSTRAPPING RESAMPLING APPROACH FOR THE SIX CONSUMER STUDIES.

Study	Average RV coefficient across simulations		
	Session 1	Session 2	Average
1*	0.967	0.970	0.969
2**	0.812	0.826	0.819
3* [a]	0.980	0.980	0.980
4* [b]	0.983	0.987	0.985
5** [a]	0.946	0.942	0.944
6** [b]	0.958	0.973	0.966

* Large differences among samples, ** Small differences among samples, [a] samples with flavor differences, [b] samples with texture and flavor differences

TABLE 2. ESTIMATION OF GLOBAL AND INDIVIDUAL REPRODUCIBILITY OF PROJECTIVE MAPPING IN THE SIX CONSUMER STUDIES, USING THE RV COEFFICIENT BETWEEN SAMPLE CONFIGURATIONS OF THE TWO EVALUATION SESSIONS.

Study	Intersession interval	Number of consumers	Product	Number of samples	Global RV coefficient between sessions	Consumer individual reproducibility (#)			
						Minimum individual RV coefficient	Maximum individual RV coefficient	Average individual RV coefficient	Percentage of consumers with significant RV coefficient ($p \leq 0.05$)
1*	2 days	91	Plain crackers	8	0.960	0.001	0.958	0.422	34%
2**	2 days	89	Plain crackers	8	0.770	0.001	0.746	0.251	15%
3* [a]	14 days	48	Vanilla milk desserts	8	0.980	0.009	0.975	0.520	54%
4* [b]	14 days	48	Vanilla milk desserts	8	0.960	0.015	0.951	0.516	50%
5 [a]**	14 days	42	Vanilla milk desserts	8	0.840	0.004	0.972	0.256	18%
6 [b]**	14 days	42	Vanilla milk desserts	8	0.920	0.003	0.968	0.321	15%

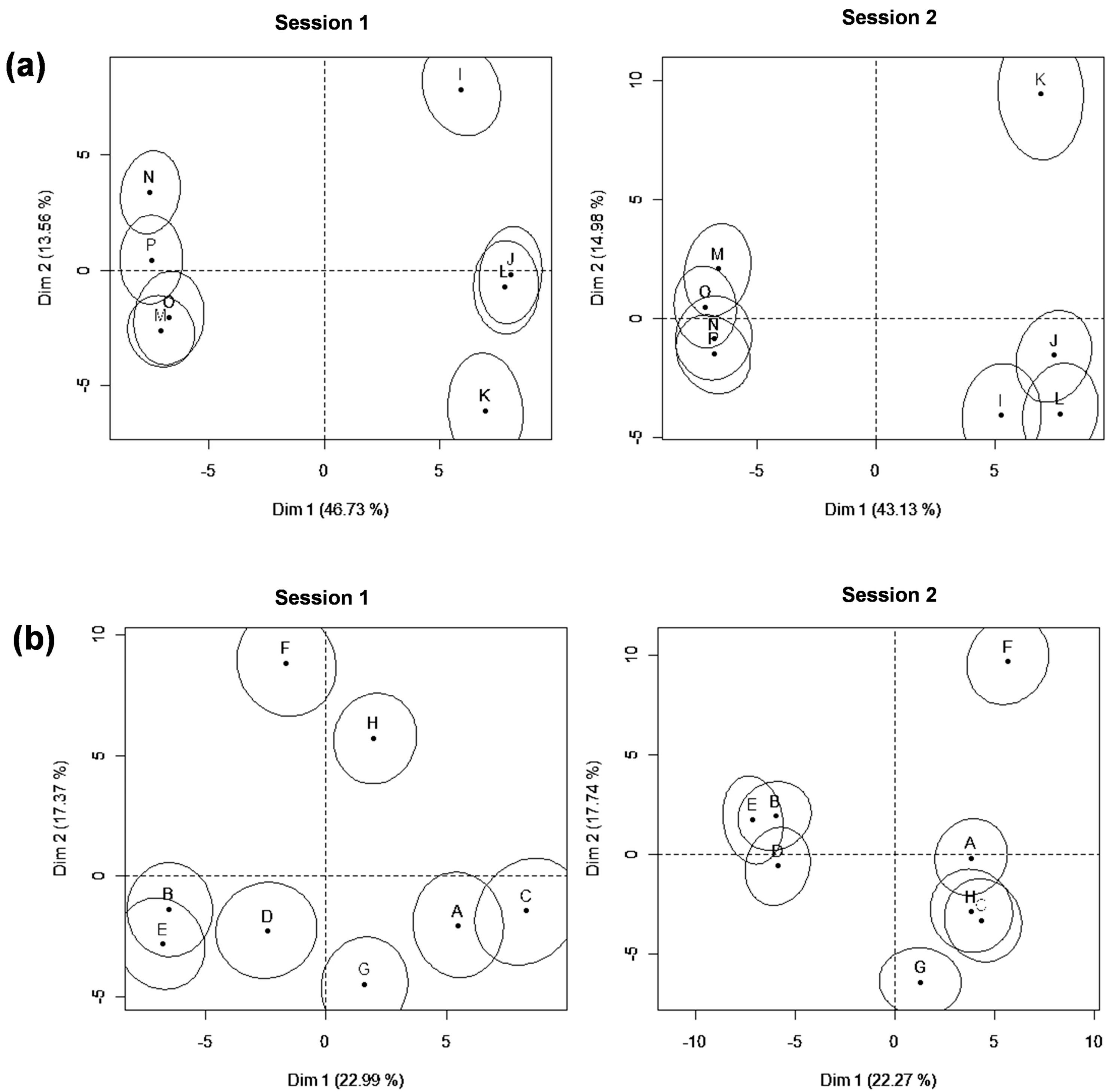
* Large differences among samples, ** Small differences among samples, [a] samples with flavor differences, [b] samples with texture and flavor differences

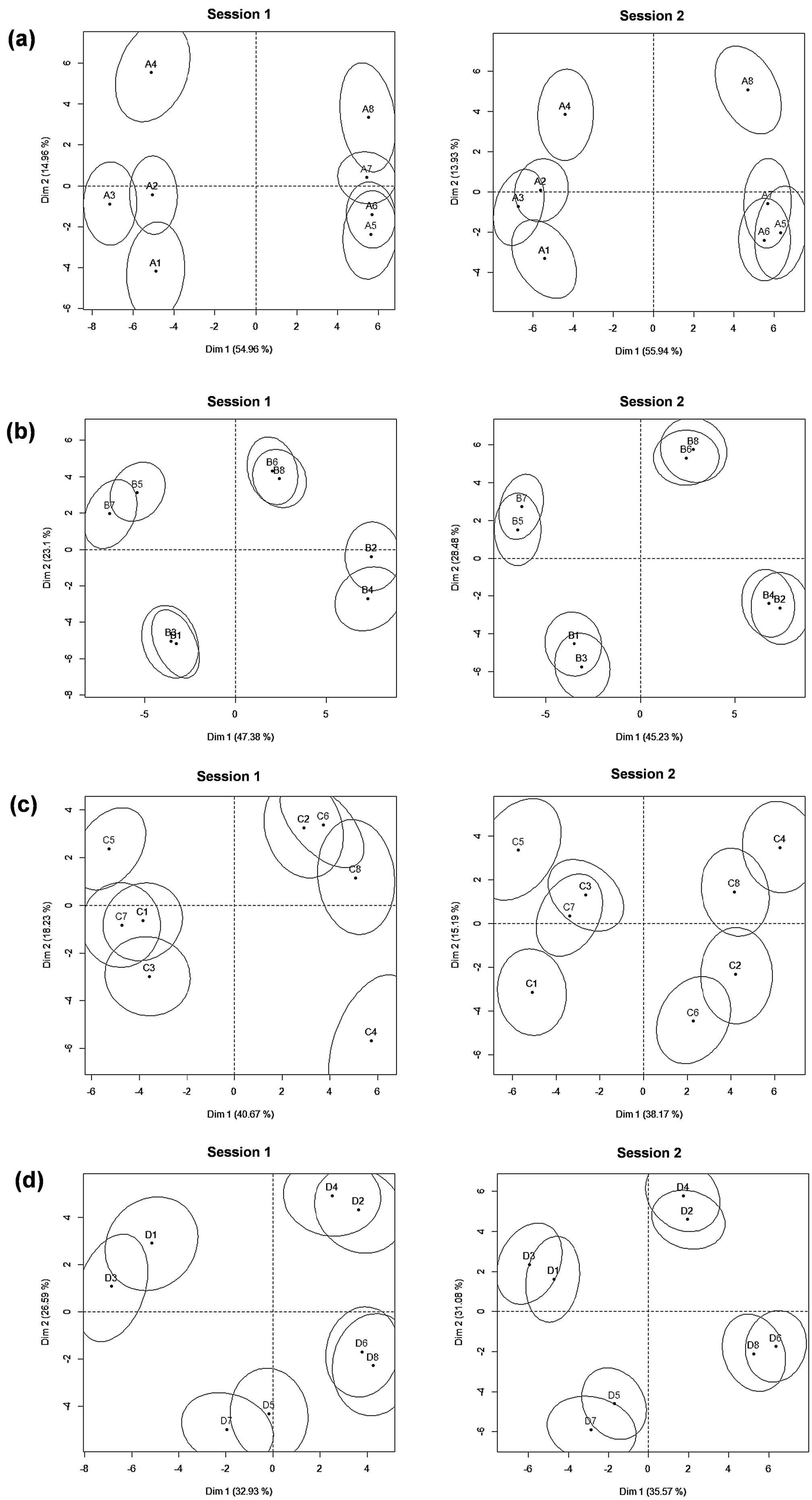
(#) Individual reproducibility was estimated using the RV coefficient between individual sample configurations between the two sessions.

TABLE 3. TOTAL NUMBER OF TERMS AND CONSENSUAL TERMS FOR THE DESCRIPTION PHASE OF PROJECTIVE MAPPING FOR THE TWO SESSIONS OF THE SIX CONSUMER STUDIES.

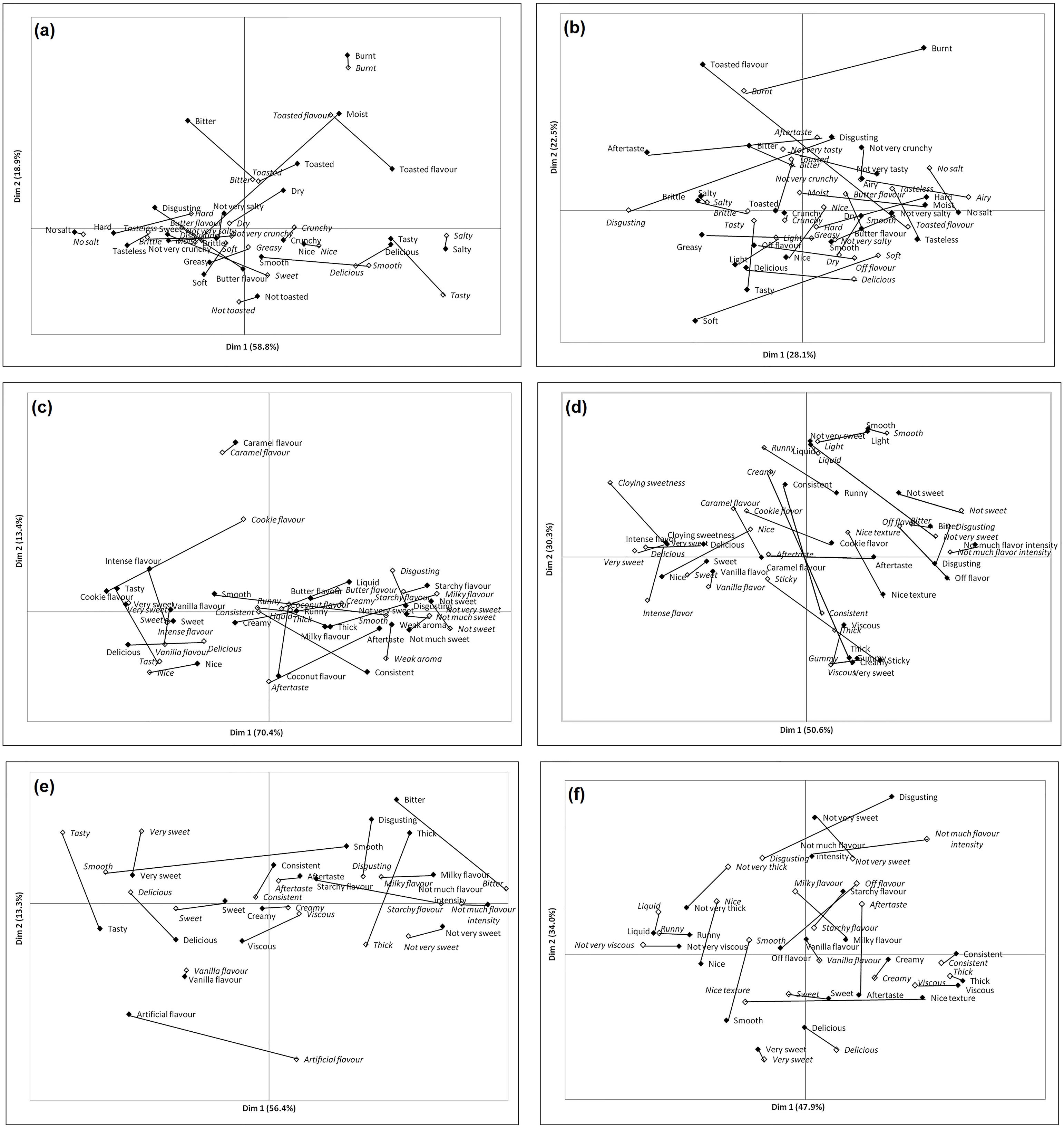
Study	Session	Total number of terms	Number of common terms between sesssions	Number of consensual terms at p≤0.10	Number of common consensual terms between sessions	RV coefficient between sessions from MFACT
1*	1	30	24	13	6	0.98
	2	26		12		
2**	1	35	27	6	2	0.80
	2	28		4		
3* [a]	1	29	25	8	6	0.98
	2	37		17		
4* [b]	1	31	27	16	12	0.94
	2	35		18		
5** [a]	1	20	18	4	0	0.81
	2	27		5		
6** [b]	1	27	22	10	8	0.94
	2	26		11		

* Large differences among samples, ** Small differences among samples, [a] samples with flavor differences, [b] samples with texture and flavor differences


(a)
Session 1
Session 2
Dim 1 (46.73 %)
Dim 2 (13.56 %)
Dim 1 (43.13 %)
Dim 2 (14.98 %)
I
N
P
J
L
O
M
K
(b)
Session 1
Session 2
Dim 1 (22.99 %)
Dim 2 (17.37 %)
Dim 1 (22.27 %)
Dim 2 (17.74 %)
F
H
B
E
D
A
C
G



(a)
Session 1
Dim 1 (54.96 %)
Dim 2 (14.96 %)
A1
A2
A3
A4
A5
A6
A7
A8
Session 2
Dim 1 (55.94 %)
Dim 2 (13.93 %)
A1
A2
A3
A4
A5
A6
A7
A8
(b)
Session 1
Dim 1 (47.38 %)
Dim 2 (23.1 %)
B1
B2
B3
B4
B5
B6
B7
B8
Session 2
Dim 1 (45.23 %)
Dim 2 (28.48 %)
B1
B2
B3
B4
B5
B6
B7
B8
(c)
Session 1
Dim 1 (40.67 %)
Dim 2 (18.23 %)
C1
C2
C3
C4
C5
C6
C7
C8
Session 2
Dim 1 (38.17 %)
Dim 2 (15.19 %)
C1
C2
C3
C4
C5
C6
C7
C8
(d)
Session 1
Dim 1 (32.93 %)
Dim 2 (26.59 %)
D1
D2
D3
D4
D5
D6
D7
D8
Session 2
Dim 1 (35.57 %)
Dim 2 (31.08 %)
D1
D2
D3
D4
D5
D6
D7
D8


(a)
Burnt
Burnt
Toasted flavour
Moist
Bitter
Toasted
Bitter
Toasted
Toasted flavour
Dry
Disgusting
Not very salty
Hard
Butter flavour
Dry
No salt
Hard
Tasteless
Sweet
Not very salty
Crunchy
No salt
Brittle
Moist
Disgusting
Not very crunchy
Tasteless
Not very crunchy
Soft
Brittle
Greasy
Crunchy
Nice
Nice
Tasty
Salty
Salty
Delicious
Greasy
Smooth
Sweet
Delicious
Smooth
Soft
Butter flavour
Not toasted
Not toasted
Tasty
Dim 1 (58.8%)
Dim 2 (18.9%)
(b)
Burnt
Toasted flavour
Burnt
Aftertaste
Disgusting
Aftertaste
Bitter
Not very tasty
Not very crunchy
Toasted
Bitter
Not very tasty
No salt
Not very crunchy
Airy
Tasteless
Brittle
Salty
Moist
Butter flavour
Hard
Airy
Salty
Nice
Moist
Disgusting
Brittle
Toasted
Crunchy
Dry
Smooth
Not very salty
No salt
Tasty
Crunchy
Hard
Greasy
Toasted flavour
Light
Butter flavour
Not very salty
Tasteless
Greasy
Off flavour
Smooth
Light
Nice
Dry
Soft
Delicious
Off flavour
Delicious
Tasty
Soft
Dim 1 (28.1%)
Dim 2 (22.5%)
(c)
Caramel flavour
Caramel flavour
Cookie flavour
Intense flavour
Tasty
Cookie flavour
Very sweet
Vanilla flavour
Very sweet
Sweet
Sweet
Intense flavour
Delicious
Vanilla flavour
Delicious
Tasty
Nice
Nice
Smooth
Runny
Consistent
Creamy
Liquid
Coconut flavour
Runny
Thick
Butter flavour
Liquid
Butter flavour
Creamy
Thick
Milky flavour
Coconut flavour
Aftertaste
Consistent
Disgusting
Starchy flavour
Milky flavour
Starchy flavour
Not sweet
Not very sweet
Not very sweet
Disgusting
Not much sweet
Smooth
Weak aroma
Not sweet
Aftertaste
Not much sweet
Weak aroma
Dim 1 (70.4%)
Dim 2 (13.4%)
(d)
Smooth
Not very sweet
Smooth
Light
Light
Runny
Liquid
Liquid
Creamy
Consistent
Cloying sweetness
Runny
Not sweet
Caramel flavor
Not sweet
Cookie flavor
Off flavor
Bitter
Disgusting
Nice
Nice texture
Bitter
Not very sweet
Intense flavor
Cloying sweetness
Very sweet
Cookie flavor
Not much flavor intensity
Delicious
Delicious
Not much flavor intensity
Very sweet
Aftertaste
Sweet
Aftertaste
Disgusting
Nice
Sweet
Vanilla flavor
Caramel flavor
Off flavor
Sticky
Vanilla flavor
Nice texture
Intense flavor
Consistent
Viscous
Thick
Thick
Gummy
Gummy
Sticky
Creamy
Viscous
Very sweet
Dim 1 (50.6%)
Dim 2 (30.3%)
(e)
Bitter
Disgusting
Tasty
Very sweet
Thick
Smooth
Smooth
Consistent
Disgusting
Very sweet
Aftertaste
Milky flavour
Aftertaste
Starchy flavour
Milky flavour
Bitter
Delicious
Consistent
Not much flavour intensity
Sweet
Creamy
Creamy
Starchy flavour
Not much flavour intensity
Sweet
Viscous
Tasty
Not very sweet
Thick
Delicious
Viscous
Not very sweet
Vanilla flavour
Vanilla flavour
Artificial flavour
Artificial flavour
Dim 1 (56.4%)
Dim 2 (13.3%)
(f)
Disgusting
Not very sweet
Not much flavour intensity
Not much flavour intensity
Disgusting
Not very sweet
Not very thick
Milky flavour
Off flavour
Starchy flavour
Nice
Liquid
Not very thick
Aftertaste
Liquid
Runny
Starchy flavour
Runny
Milky flavour
Smooth
Not very viscous
Not very viscous
Vanilla flavour
Off flavour
Vanilla flavour
Nice
Consistent
Creamy
Consistent
Creamy
Thick
Viscous
Thick
Nice texture
Viscous
Sweet
Sweet
Aftertaste
Nice texture
Smooth
Delicious
Very sweet
Delicious
Very sweet
Dim 1 (47.9%)
Dim 2 (34.0%)