Published in final edited form as:

Food Qual Prefer. 2015 October 1; 45: 41–49. doi:10.1016/j.foodqual.2015.05.004.

Check-All-That-Apply (CATA), Sorting, and Polarized Sensory Positioning (PSP) with Astringent Stimuli

Erin E. Fleming^{1,2}, Gregory R. Ziegler², and John E. Hayes^{1,2,*}

¹Sensory Evaluation Center, College of Agricultural Sciences, The Pennsylvania State University, University Park, Pennsylvania, USA

²Department of Food Science, College of Agricultural Sciences, The Pennsylvania State University, University Park, Pennsylvania, USA

Abstract

Multiple rapid sensory profiling techniques have been developed as more efficient alternatives to traditional sensory descriptive analysis. Here, we compare the results of three rapid sensory profiling techniques – check-all-that-apply (CATA), sorting, and polarized sensory positioning (PSP) – using a diverse range of astringent stimuli. These rapid methods differ in their theoretical basis, implementation, and data analyses, and the relative advantages and limitations are largely unexplored. Additionally, we were interested in using these methods to compare varied astringent stimuli, as these compounds are difficult to characterize using traditional descriptive analysis due to high fatigue and potential carry-over. In the CATA experiment, subjects (n=41) were asked to rate the overall intensity of each stimulus as well as to endorse any relevant terms (from a list of 13) which characterized the sample. In the sorting experiment, subjects (n=30) assigned intensity-matched stimuli into groups 1-on-1 with the experimenter. In the PSP experiment, (n=41) subjects first sampled and took notes on three blind references ('poles') before rating each stimulus for its similarity to each of the 3 poles. Two-dimensional perceptual maps from correspondence analysis (CATA), multidimensional scaling (sorting), and multiple factor analysis (PSP) were remarkably similar, with normalized RV coefficients indicating significantly similar plots, regardless of method. Agglomerative hierarchical clustering of all data sets using Ward's minimum variance as the linkage criteria showed the clusters of astringent stimuli were approximately based on the respective class of astringent agent. Based on the descriptive CATA data, it appears these differences may be due to the presence of side tastes such as bitterness and sourness, rather than astringent sub-qualities per se. Although all three methods are considered 'rapid,' our prior experience with sorting suggests it is best performed 1:1 with the experimenter, which makes sorting relatively less efficient than CATA or PSP. Based on the evaluation criteria used here, the choice of method depends on the time constraints of the experimenter and the need for descriptive terms to understand the sensory space of the samples. Accordingly, we recommend a mixed

© 2015 Published by Elsevier Ltd.

*Correspondence: Dr. John E. Hayes, Department of Food Science, Pennsylvania State University, 220 Rodney A. Erickson Food Science Building, University Park, PA, 16802, USA, jeh40@psu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

approach that combines CATA with a subsequent PSP task so that the product space can be well characterized before choosing poles for PSP.

Keywords

Rapid Profiling; Rapid Methods; Perceptual Maps; Polyphenols; Organic Acids; Salts; Astringency;

1. Introduction

Methods used to describe the sensory attributes of a product have been commonly used in the food industry since the mid 1960s. Classical descriptive analysis is one of the most powerful and widely used methods because it allows quantitative comparisons to be made across different products on specific attributes. Several types of descriptive analysis methods are commonly used in industry, as well as in academia, including the Flavor Profile Method, the Texture Profile Method, Quantitative Descriptive Analysis® (QDA®) and Spectrum™ Descriptive Analysis (Murray et al., 2001). One common limitation of these techniques is the amount of time that must be invested to recruit, train, and maintain a group of highly skilled assessors (Chambers and Wolf, 1996). As a result, there is a strong interest in the development of new rapid methods which seek to increase the efficiency of the data collection process while maintaining the robustness of the information obtained.

Here, three rapid methods – check-all-that-apply (CATA), sorting, and polarized sensory positioning (PSP) – were compared using a diverse range of astringent stimuli. These rapid methods differ in their theoretical basis, implementation, and data analysis, and the relative advantages and limitations are largely unexplored. Moreover, it is unclear how these methods may be best combined to better characterize the sensory space of a set of products.

In a CATA task, participants are given a list of terms and asked to endorse those that characterize the sample (e.g., Bennett and Hayes, 2012; Sinopoli and Lawless, 2012). These questions are often used in market research to reduce the response burden of participants (Rasinski et al., 1994), as participants do not have to generate their own descriptors for each sample. One important consideration of CATA is which words or phrases are included by the experimenter in the list, as excluding salient characteristics may result in data that does not accurately reflect important or relevant product attributes. The generation of the terms used in CATA can be done in a number of ways: by a trained panel, by consumers during the test (modified free choice profiling), or by simple term generation during focus groups (Dooley et al., 2010). Likewise, terms can be used from previous studies on similar products.

Historically, sorting has been used extensively in psychology, anthropology, and sociology for data collection (Miller et al., 1986); in a sorting task, the participant examines a number of objects and systematically groups them into some number of categories based on their own criteria (Coxon, 1999). For sensory applications, sorting typically estimates the global degree of similarity between stimuli across a group of individuals. Untrained participants may find sorting a large number of stimuli into groups difficult, especially when asked to

complete the task without guidance from the experimenter. Accordingly, attribute lists are sometimes provided to assessors to help guide the grouping process (Lelièvre et al., 2008). The number of assessors required for a sorting task depends on the amount of training. Typically, when using trained assessors, the task requires between 9–15 individuals (Cartier et al., 2006), compared to 9–98 individuals with untrained assessors/naïve consumers (Cadoret et al., 2009).

Polarized Sensory Positioning (PSP) is a recently described method which involves rating the relative similarity/dissimilarity of all products compared to a number of fixed references, which are termed ‘poles’. This methodology was originally developed to compare the sensory characteristics of water by Teillet and colleagues (2010). In a PSP task, participants are asked to rate each sample’s similarity/dissimilarity relative to each pole. From the perspective of the participant, this is very similar to a classical Difference From Control test, except that the test stimulus is being compared to multiple poles, rather than to a single reference. The main advantage of PSP (assuming that the same poles are used between sessions) is that comparisons can be made across multiple sessions or sample sets. In other words, as long as identical poles are used for each session, data from multiple sessions even over weeks or months can be aggregated. A key step of this method is the selection of the poles. When selecting poles it is critical that these reference samples cover the whole sensory space (de Saldamando et al., 2013). Typically, 3 poles have been used in a PSP task (Teillet et al., 2010; de Saldamando et al., 2013; Cadena et al., 2014), although in theory more can be used.

Oral astringency can be defined as a “complex of sensations due to shrinking, drawing, or puckering of the epithelium as a result of exposure to substances such as alums or tannins” (ASTM, 1991). Astringency collectively describes a group of sensations that include drying, puckering, velvety, and roughing, which are typically elicited by stimuli like polyphenols (e.g. tannins), multivalent salts (e.g. alum), organic acids (e.g. malic acid), dehydrating agents (e.g. ethanol), and charged polysaccharides (e.g. chitosan) (Joslyn and Goldstein, 1964; Lawless et al., 1996; Kallithraka et al., 1997; Peleg et al., 1998; Rodriguez et al., 2003).

Sampling a large number of astringent compounds in one session presents some methodological difficulty, as astringent compounds exhibit build-up or ‘carry-over’ upon repeated exposure. Berg et al. (1955) first noted this effect and showed that small differences in tannin concentration were difficult to determine in paired tests due to this phenomenon. Theoretically, the high carry-over of astringent compounds would present a challenge in classical descriptive analysis: both during training sessions (i.e. tasting multiple references) as well as test sessions where a large number of samples would be presented. Here, we reasoned that comparing a broad range of astringent compounds susceptible to carry-over — and therefore problematic if compared using classical descriptive analysis techniques — would serve as a unique opportunity to both elucidate additional information about the diverse sensations associated with astringent stimuli, as well as to compare three novel rapid methods. The aim of the present study was to compare the utility of three rapid profiling methods – CATA, sorting, and PSP – in characterizing a broad range of astringent stimuli. Here, our criteria for ‘utility’ were adapted from the approach described by

Reinbach and colleagues (2014). The criteria were: a) Configurational Agreement, b) Descriptive ability, and c) Practicality.

2. Materials and Methods

2.1. Participants

A total of 112 reportedly healthy individuals were recruited from the Pennsylvania State University campus and surrounding area (State College, PA) via email after prescreening for eligibility and their willingness to participate. Approximately one fourth of the participants took part in the sorting task (n=30) and the remainder participated in either the CATA task (n=41) or PSP (n=41). Criteria for eligibility included: between 18–64 years old; not pregnant or breastfeeding; no known defects of smell or taste; no lip, cheek, or tongue piercings; nonsmoker (had not smoked in last 30 days); no food allergies or sensitivities; no history of choking or difficulty swallowing. Participants provided informed consent and were paid for their time (45–60 min depending on specific session). All procedures were approved by the Pennsylvania State University Institutional Review Board.

2.2. Stimuli

Ten representative food grade astringent agents were chosen from three broad classes of astringent compounds: multivalent salts (MS), organic acids (OA), and polyphenols (P) (Table 1). These stimuli were served as aqueous solutions made with reverse osmosis (RO) water. Appropriate concentrations for each sample were determined based on an intensity matching experiment that was incorporated into the CATA experiment. As a test for reliability across experiments, a duplicate ammonium aluminum sulfate ('alum') stimulus was included in the sample set (Falahee and MacRae, 1997; Lim and Lawless, 2005).

Across the three experiments, participants were instructed to securely fasten a plastic nose clip (Cressi, Genoa, Italy) to the outside of their noses prior to receiving their first sample. This was done to minimize the olfactory (both ortho- and retro-nasal) contribution to the participants' perceptual judgments, since many of the stimuli exhibit strong, characteristic aromas. Given our interest in exploring the sub-qualities of each class of astringent agent, we reasoned that using a nose clip would aid the participants in focusing on taste and touch sensations rather than overall flavor.

All stimuli were presented at room temperature in 10 mL aliquots. Sampling procedures were kept consistent between studies and were as follows: "Please pour the entire contents of sample XXX into your mouth and swish the solution for 5 seconds as if using mouthwash. Spit the sample into the covered spit cup provided after the 5 seconds." Presentation order was counterbalanced across participants using a Williams design. Participants rinsed with room temperature RO water prior to the first stimulus and between each subsequent stimulus. A minimum interstimulus interval (ISI) of 2 min was enforced between samples via software. All tests were completed at the Sensory Evaluation Center in the Department of Food Science at the Pennsylvania State University. Compusense five software, version 5.2 (Guelph, Ontario, Canada) was used for data collection in the CATA

and PSP experiments. Websort (Optimal Product Ltd) was used for data collection in the sorting experiment.

2.3. CATA

As part of the intensity matching study used to identify doses for the sorting and PSP experiments, subjects (n=41) were asked to endorse any number of terms from a list of 13 descriptors, shown in Table 2. The terms were generated from previous studies examining the time-course of astringent compounds (Lee and Lawless, 1991); basic tastes and other chemesthetic sensations were also included. While acknowledging that naïve consumers may have difficulty distinguishing between some of the terms found in the CATA experiment, we did not provide references or lengthy definitions as the goal here was to compare rapid methods which require little or no training. For additional information on astringent descriptors and references, see works by Gawel and colleagues (2000; 2001). To minimize dumping, an option for 'Other' was included with an open-ended text box; these data were not used in any analysis. Presentation order of the CATA terms on the computer screen was counterbalanced within and across participants using a Williams design to avoid order biases. All evaluations were conducted in isolated sensory booths under normal white lighting.

Appropriate doses for the subsequent sorting and PSP experiments were determined based on an intensity matching experiment which was part of the CATA task. Participants were asked to rate the overall intensity of each stimulus using a 10 cm unstructured line scale ranging from 'weak' to 'strong'. Average overall intensity ratings falling within 4.5 $\pm$ 1.5 were considered approximately equal (data not shown).

2.4. Sorting

All sorting data were collected one-on-one in a windowless clinical-style examination room under normal white lighting, with the experimenter seated across the table from the participant. The participants (n=30) were instructed to taste the samples and to note any tastes and sensations they experienced. A pen and notepad were provided for note taking along with a list of descriptors identical to that used in the CATA task; participants were reminded that this list was non-exhaustive and they could use their own terms as a basis for the groups as needed. Participants were instructed to make a minimum of two groups and a maximum of 10 groups with no other constraints.

Typically, 'sorting' refers to a 'free sorting' task in which there are no constraints imposed on the assessor, who simply divides the stimuli into a number of groups based on whatever criteria he or she deems most appropriate. However, a number of variations on sorting exist. One such variant, 'directed sorting,' consists in providing information on either the number of groups and/or the criteria for group formation (Valentin et al., 2012). Here, consistent with free sorting, no sorting criteria were provided, and participants were allowed to use as few or as many groups as they deemed appropriate, as long as all samples were not in one giant group, or in 11 groups of one item each. However, since a non-exhaustive list of terms (Table 2) was provided to participants, the task was not completely semantic free; thus, it may be more appropriate to describe the present task as 'modified free sorting'. Critically,

the experimenter emphasized that the list of terms provided was merely a starting point and not a comprehensive list, so we do not believe this modification substantially alters the validity of our results.

2.5. Polarized Sensory Positioning

Participants (n=41) first received a tray with the three reference samples ('poles') as well as a pen and notepad for note taking; the 3 poles (alum, tannic acid, and malic acid) were selected to represent each class of astringent agent: multivalent salt, polyphenol, and organic acid, respectively. After tasting each reference sample, they were asked to note the taste(s) or sensation(s) they experienced. Participants kept these notes and were able to reference them during the remainder of the test. Reference samples were collected from participants prior to their receiving any additional stimuli (i.e. re-tasting was not allowed).

Participants rated each unknown stimulus relative to the three poles on a 10 cm unstructured line scale ranging from 'exactly the same'=0 to 'completely different'=10. Blind duplicates of the three poles as well as a duplicate alum sample were included within the stimulus set. Instructions for ratings were as follows: 'Please remember to make your ratings based on the similarity/dissimilarity of the tastes and sensations and not based on how the samples appear.' All stimuli were tasted in isolated testing booths under normal white lighting.

2.6. Data Analysis

All statistical analyses were performed using R version 2.14.1 (R Project for Statistical Computing). FactoMineR (Lê et al., 2008a) was used to perform correspondence analysis (CA), multiple factor analysis (MFA) and to calculate NRV coefficients. MDS and Cochran's Q test were carried out using the 'Smacof' (De Leeuw and Mair, 2011) and 'RVAideMemoire' (Abdi et al., 2013) packages, respectively. Additional details are provided below.

2.6.1. Check-all-that-apply (CATA)—For each stimulus, a data matrix was created containing the CATA term in columns and participants in rows with each cell indicating if the term was checked or not (1/0, respectively). Cochran's Q test (Meyners et al., 2013) was used to determine significant differences among stimuli for each descriptor (Table 4). This nonparametric statistic is appropriate for the analysis of two-way randomized block designs to determine whether k treatments (i.e. samples in the test) have identical effects when the response variable (i.e. CATA term) is binary (Varela and Ares, 2012).

Frequency counts for each of the terms were tabulated across stimuli and converted to a percentage. These data were organized into a stimulus x CATA term matrix containing the percentage of participants who checked each CATA term for each stimulus (Table 4). Correspondence Analysis (CA) was carried out on this matrix to determine the spatial configuration of the stimuli as well as the CATA terms. CA has been used previously to analyze frequency counts obtained through CATA questions in order to obtain a bi-dimensional representation of the stimuli and CATA terms (Ares et al., 2010; Varela and Ares, 2012).

For cluster analysis, the coordinates (two dimensions) obtained from CA were submitted to agglomerative hierarchical clustering with Ward's minimum variance method for the linkage criteria. Agglomerative hierarchical clustering begins with each observation in its own cluster, then merges the clusters one-by-one based on proximity until only one large cluster containing all observations remains (Lawless, 2013). As for the linkage criteria, Ward's minimum variance minimizes the increase in total within-cluster sum of squared error, thereby incorporating cluster homogeneity and cluster separability into the cluster criterion (Murtagh, 1985).

2.6.2. Sorting—Sorting data is most commonly analyzed using multidimensional scaling (MDS), a technique which aims to create a spatial map representing the relative similarities and differences between stimuli (Lawless et al., 1995; Varela and Ares, 2012). The number of times each sample was placed in the same group was summed across the 30 participants to create an index of similarity. A dissimilarity matrix was then determined followed by dimension and stress value analysis. In order to determine the appropriate number of dimensions, Kruskal's stress was calculated. Based on the scree plot generated from these stress values (data not shown), a two-dimensional MDS solution was chosen. Agglomerative hierarchical cluster analysis was performed on a dissimilarity matrix of the group data using Ward's minimum variance as the linkage criteria.

2.6.3. Polarized Sensory Positioning (PSP)—Data from PSP were analyzed using MFA to preserve individual data and to compensate for different responses across individuals. The data matrix consisted of individual ratings for each stimulus across each of the three poles, as done previously (Ares et al., 2013; de Saldamando et al., 2013; Cadena et al., 2014). In other words, MFA was performed considering data from each participant as a separate group of variables. MFA is a multivariate technique which aims to integrate different groups of variables describing the same observations (Reinbach et al., 2014). It can be used to analyze the similarity of a set of observations explained by several groups of variables on comparable or contradictory scales (Abdi et al., 2013), as it is able to balance the influence of each group of variables to demonstrate patterns of attribute correlation (Lê et al., 2008b; Nestrud and Lawless, 2008).

To account for the poles, duplicates of each pole were included in the sample set as done previously (Cadena et al., 2014). Alternatively, poles can be accounted for in the analysis by adding supplementary rows of synthetic data for each pole consisting of a distance of '0' from itself and a distance of '10' to the other poles (Teillet et al., 2010). Here, data were analyzed using both approaches, with the latter producing severely compressed visual configurations with poor resolution; therefore, the duplicate pole approach of Cadena and colleagues was used. Agglomerative hierarchical cluster analysis was performed on the coordinates from the two-dimensional solution obtained from MFA using Ward's minimum variance as the linkage criteria.

2.6.4. Composite MFA—Each of the two-dimensional solutions from the three methods was analyzed using MFA in order to generate a composite MFA perceptual map. That is, the coordinates generated from CA, MDS, and MFA from CATA, sorting, and PSP,

respectively, were treated as individual MFA groups for analysis, thus allowing comparison between this composite configuration and each individual method's perceptual map.

3. Results

3.1. Configurational Congruence

The criteria for configurational congruence was assessed in three ways: a) by the visual plots generated by CA, MDS, and MFA; b) by the NRV coefficients comparing these plots for significance; and c) by the dendrograms generated via agglomerative hierarchical clustering.

Figure 1 shows the CA plot for CATA (A), the MDS plot for sorting (B), the MFA plot for PSP (C) and the composite MFA plot (D) generated from the three individual two dimensional solutions from each method. In the CATA map from CA (Figure 1D), words which were endorsed by less than 25% of participants are shown in gray. CA, MDS, and MFA belong to a family of multivariate statistical techniques that are commonly used to represent the similarity of items within a data set. Namely, these methods can be used to visually represent the magnitude of perceptual distances between stimuli and can provide information on how participants group these samples. The location of stimuli (or terms) in these perceptual maps can be interpreted as a measure of their similarity to one another. Generally, these plots illustrate the tendency for organic acids, polyphenols (and zinc chloride), and the alum duplicates to group together on the left, right, and central portions of the plots, respectively.

Normalized RV coefficients (NRV) were used to compare sample configurations obtained for each method (Table 3). NRV coefficients were calculated for all possible combination of methodologies for two dimensional solutions and were used to compare the similarity of the perceptual maps. The NRV is an approximation to a permutation test for significance of RV coefficients, and is preferred here because the magnitude of an RV depends on the number of stimuli and dimensions between plots. The NRV coefficient is interpreted in the same way as a z-score for a normal distribution, with a large score (>2) indicating significant similarity between configurations (Nestrud and Lawless, 2010).

The dendrograms for the agglomerative-linkage cluster analysis using Ward's minimum variance as the linkage criteria for CATA (A), sorting (B), and PSP (C) are shown in Figure 2. Generally, the cluster analysis indicates that astringent compounds grouped based on class (i.e. multivalent salt, organic acid, or polyphenol) across each of the experiments. For CATA (Figure 2A) the organic acids clustered first with the multivalent salts, before finally forming their own distinct cluster. The same trend was observed the PSP (Figure 2C). For sorting (Figure 2B), the polyphenols clustered first with the alum duplicates, before forming their own cluster. In all dendrograms, zinc chloride (a multivalent salt) clustered with the polyphenols rather than with the alum duplicates.

3.2. Descriptive Ability

Because both sorting and PSP lack semantic information, descriptive data was analyzed for the CATA experiment. Table 4 shows the percentage of participants who endorsed each term across stimuli. Significant differences across samples for individual attributes were

observed for all terms except 'umami/savory' and 'tickle.' Descriptive data can also be found in Figure 1A, as the CATA terms were included in the CA plot.

3.3. Practical Considerations

Although CATA, sorting, and PSP are all considered rapid methods as compared to classic descriptive analysis, these experiments differed greatly in their practicality; namely, the time burden on the experimenter. For example, despite the fact that the number of participants for all experiments were kept relatively consistent, because sorting (n=30) was conducted 1-on-1 with the experimenter, collection of sorting data took nearly eight times longer relative to CATA (n=41) and PSP (n=41), which were collected 12 at time in isolated sensory booths. These considerations are discussed further below.

4. Discussion

In the present study, the relative utility of three rapid sensory methods were compared using three specific criteria: configurational congruence, descriptive ability, and practical considerations. These methods were carried out using a diverse range of astringent stimuli including multivalent salts, organic acids, and polyphenols.

The three test methods provided comparable visual configurations in spite of dramatically different task demands on the subjects, analytical approaches (CA, MDS, and MFA for CATA, sorting and PSP, respectively), and time demands on the experimenter. Upon visual inspection of the plots (Figure 1), the three methods resulted in similar configurations with organic acids, multivalent salts, and polyphenols grouping relatively close to one another with the exception of zinc chloride, which grouped more closely with polyphenols than multivalent salts, and this similarity is confirmed by the NRV coefficients.

The dendrograms (Figure 2) show that astringent compounds clustered approximately on the class of astringent in all three experiments. Also, alum clustered with its blind duplicate across all methods, suggesting good reliability. Generally, the polyphenols tended to cluster together, as did the organic acids. However, zinc chloride tended to cluster with the polyphenols, rather than alum, across all three experiments. While we expected zinc chloride to group with alum and its blind duplicate across all plots, this compound consistently grouped with the polyphenols in all 3 perceptual maps (Figure 1) and in all dendrograms (Figure 2). This cluster was inconsistent with the results of Lim and Lawless (2005) who observed that alum and zinc chloride clustered together in an 'astringent group', as opposed to 'bitter', 'metallic', or 'salty' groups. However, based on CATA results here (see Table 4), we believe this is likely due to a similar endorsement of the terms '*sour*', '*stinging/pricking*', and '*other*' between zinc chloride and the polyphenols. Likewise, key differences between the alum and zinc chloride include greater '*salty*' and lower '*sour*' and '*sweet*' endorsements for zinc chloride, which were more consistent with CATA endorsements for polyphenols than the alum duplicates.

Here, tannic acid consistently clustered with other polyphenols rather than with alum. This pattern was inconsistent with previous studies showing that tannic acid clustered with alum when characterized using pair-wise comparisons and MDS (Kielhorn and Thorngate, 1999).

However, dramatically different sets of stimuli were used here as compared to the Kielhorn and Thorngate study, suggesting context effects likely influenced the groups. In other words, when presented with a broader set of stimuli intended to represent a wide range of tastes and oral sensations, tannic acid is perceived as more similar to alum; here, when compared amongst a set of exclusively astringent stimuli, tannic acid is more like the polyphenols. Similar contextual effects have been seen previously with odors (Lawless et al., 1991). Another possibility for this result was the marked differences in 'sweet' and 'sour' ratings between alum and tannic acid (Table 4). In addition to astringency, alum has previously been shown to elicit both sweet and sour sensations (Breslin et al., 1993). Here, 'sweet' and 'sour' were each endorsed by greater than 25% of participants (see Table 4) for alum while they were only endorsed 0% and 14.6% of the time, respectively, for tannic acid. While it was unexpected that alum would exhibit such a high percentage of endorsements for the term 'bitter', given the use of naïve consumers in this task, and the common confusion with bitterness and astringency (or 'drying') (Green, 1993), this is not surprising.

The RV coefficient (Robert and Escoufier, 1976) is a multivariate generalization of Pearson's R^2 , and is commonly used to measure the similarity between multivariate configurations (Ares et al., 2013; Reinbach et al., 2014). Here, the normalized RV coefficient (NRV) was used due to potential differences in the number of stimuli in a group and number of dimensions in the perceptual maps (Nestrud and Lawless, 2008). The NRV can be interpreted similarly to a z-score, with a large score (>2) indicating significant similarity between the maps.

The NRV coefficients calculated between each pair of methods show significantly similar plots, regardless of method (p 's <0.001). Similarly, Reinbach and colleagues recently compared CATA, CATA with intensity, and Napping with eight beers, and found that both qualitative and quantitative analysis revealed a high agreement between the three methods in terms of perceived product differences (Reinbach et al., 2014). These results also showed that the precision and reproducibility of sensory information obtained by consumers via CATA is comparable to that of Napping. Likewise, in a study comparing CATA, Napping, and PSP with functional yogurts, Cadena and others (2014) found that all three methods provided similar information when using untrained assessors. Based on these findings, it is not unreasonable that significantly similar plots were found here when comparing sorting, CATA, and PSP.

Regarding the qualitative aspects of the stimuli, the results of Cochran's Q test indicate that the CATA experiment was able to distinguish differences between astringent compounds, as shown in Table 4. Generally, 'drying' and 'bitter' were endorsed most frequently (i.e. greater than 35% across all stimuli for each respective astringent class), for the multivalent salts and polyphenols versus 'drying', 'sour', and 'puckering' for the organic acids. This is reflected in the perceptual map generated from CATA using CA (Figure 1A) which shows both the visual configurations between stimuli and CATA descriptors. Here, the acids grouped near 'sour' and 'puckering' whereas the polyphenols grouped near 'bitter,' 'roughing,' and 'drying.'

In addition to sourness, organic acids have been shown to elicit astringent sensations (Thomas and Lawless, 1995; Lawless et al., 1996; Sowalsky and Noble, 1998). Moreover, Rubico and McDaniel (1992) found that free-choice profiling of organic and inorganic acids analyzed by generalized Procrustes Analysis (GPA) resulted in three principal axes; the first of which was astringency/mouthfeel while bitterness and sourness were the most important for the second and the third principal axes, respectively. Therefore, the characterization of organic acids as mostly 'drying' and 'puckering' was expected, as these are well-documented astringent sub-qualities.

Among the four terms often associated with astringency in the CATA term list ('drying', 'roughing', 'puckering', and 'velvety'), 'roughing' and 'velvety' were endorsed relatively less often (i.e. less than 40% and 30%, respectively, across all stimuli). Work by Lee and Lawless (1991) showed that the astringent sub qualities 'drying', 'roughing', and 'puckering' de-correlated over time and were not totally interchangeable; therefore, it is not surprising that these terms were not consistent between all astringent compounds. The only terms in which significant differences were not detected across stimuli were 'umami/savory,' and 'tickle.' Because these stimuli were chosen to represent broad classes of astringent compounds, it is not unexpected that these did not significantly differ on these attributes, as they are not commonly associated with astringent sensations.

Although CATA, sorting, and PSP are all considered 'rapid' methods as compared to classic descriptive analysis, the practical considerations of conducting these types of experiments cannot be overlooked. For example, both CATA and PSP were conducted in individual sensory booths, whereas sorting was carried out 1-on-1 with the experimenter. Previously, a pilot sorting study in our laboratory was conducted in which a sorting task was completed in individual sensory booths. Based on the feedback from participants and the lack of consistency in the data, we felt the sorting task was too cognitively difficult for untrained participants to accurately complete while in individual sensory booths. Therefore, the decision was made to complete this study 1-on-1 with the experimenter.

In terms of simple practicality, the PSP and CATA methods appear to be more time efficient than sorting. While sorting is considered a rapid method compared to traditional pairwise comparisons (e.g. 55 sample pairs would be needed for the 11 stimuli tested here), or to traditional descriptive analysis methods, it is unclear if this method can be conducted by untrained assessors without the oversight of the experimenter. Comparatively, PSP and CATA are much more efficient, as multiple subjects can be tested in parallel within isolated sensory booths rather than one-on-one with the experimenter. In our case, this reduced total experimenter time from 30 hours for sorting to 4 hours each for PSP and CATA. However, this conclusion may not hold for trained assessors (as opposed to naïve consumers), assuming they can successfully compete a sorting task without the direct involvement of the experimenter.

Although numerous works comparing the results and usefulness of rapid methods have been reported recently (King et al., 1998; Dehlholm et al., 2012; Cadena et al., 2014; Reinbach et al., 2014), this is the first that specifically compares CATA, sorting, and PSP. Moreover, this

is the first focused on astringent stimuli, as these compounds are extremely difficult to characterize using classic descriptive analysis due to high fatigue and carry-over.

As their name implies, rapid sensory methods have the ability of being completed more quickly than traditional descriptive analysis. However, unlike descriptive analysis, these methods do not include direct scaling of multiple attributes, which may be extremely helpful for product developers when reformulating products. Also, the most appropriate method for exploratory research may not generalize to an industrial setting focused on specific products, where intensity ratings may be more valuable.

Here, the relative advantages and limitations of CATA, sorting, and PSP were explored. Based on these findings, the most important step in designing a CATA experiment is deciding which terms to include in the ballot. For PSP, an understanding of the sensory profiles of the stimuli is key so that appropriate poles that represent the sensory space can be chosen. For sorting, it is unclear if this task can be carried out by naïve consumers without the supervision of the experimenter; particular care has to be taken to ensure that participants are capable of understanding and completing the task.

5. Conclusion

Here, three broad classes of astringent stimuli were compared using three rapid sensory methods. Astringent stimuli tended to group together based on class (i.e. multivalent salt, organic acid, or polyphenol) consistently between plots as well as in dendrograms generated via hierarchical cluster analysis across all methods. Generally, ‘drying’ and ‘bitter’ were endorsed most frequently (i.e. greater than 35% across all stimuli for each respective astringent class), for the multivalent salts and polyphenols versus ‘drying’, ‘sour’ and ‘puckering’ for the organic acids. Collectively, the CATA endorsements and clustering data suggest additional work is needed to determine if these patterns are driven by differences in astringent sub-qualities or by other side tastes such as bitterness or sourness. In the future, citric acid and quinine could be included as additional poles in a PSP task to determine the relative importance of side tastes versus the various astringent sub-qualities.

Regarding the rapid methods used here, these data suggest CATA, sorting, and PSP are all viable methods for the rapid sensory characterization of astringent stimuli. The perceptual maps across methods were visually similar, and were significantly similar when tested with the NRV coefficient. The composite map across all three methods (Figure 1D) shows that the stimuli grouped generally by class, with the exception of zinc chloride, which fell near the polyphenols. Based on the criteria we selected a priori to compare methods (i.e. configurational congruence, descriptive ability, and practical considerations), we recommend that CATA and PSP be used sequentially, with CATA being used to explore and characterize the sensory space and guide the selection of the poles, followed by PSP for additional refinement.

Acknowledgments

This manuscript was completed in partial fulfillment of the requirements for a Master of Science degree at the Pennsylvania State University by the first author. The authors would like to thank Rachel Antenucci MS, Rachel

Primrose MS, Alissa A. Nolden MS, and Drs. Nadia Byrnes, Emma Feeney, and Toral Zaveri, for assistance in data collection and protocol development, as well as Dr. Michael Nestrud, for his advice on data analysis. We also thank the Pennsylvania State University Statistical Consulting Center for additional guidance as well as our study participants for their time and participation.

Conflict of interest disclosure

EEF is supported by the Ralph Lee Graduate Fellowship Award at the Pennsylvania State University; this award is underwritten by the Pennsylvania Manufacturing Confectioners' Association (PMCA). JEH and GRZ have each received speaking or consulting fees from corporate clients in the food industry. Additionally, the Sensory Evaluation Center at Penn State routinely conducts taste tests for the food industry to facilitate experiential learning for students. None of these organizations have had any role in study conception, design or interpretation, or the decision to publish these data. JEH also receives support from a National Institutes of Health grant from the National Institute of Deafness and Communication Disorders [DC010904].

References

- Abdi H, Williams LJ, Valentin D. Multiple factor analysis: principal component analysis for multitable and multiblock data sets. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2013; 5:149–179.
- Ares G, Barreiro C, Deliza R, Giménez A, Gambaro A. Application of a check-all-that-apply question to the development of chocolate milk desserts. *Journal of Sensory Studies*. 2010; 25:67–86.
- Ares G, et al. Polarized projective mapping: Comparison with polarized sensory positioning approaches. *Food Quality and Preference*. 2013; 28:510–518.
- ASTM. Standard terminology relating to sensory evaluation of materials and products. Annual book of ASTM Standards 15.07-end use products. Philadelphia, PA: American Society for Testing Materials; 1991. p. 1-3.E 253–91a
- Bennett SM, Hayes JE. Differences in the chemesthetic subqualities of capsaicin, ibuprofen, and olive oil. *Chemical senses*. 2012; 37:471–478. [PubMed: 22281531]
- Berg H, Filippello F, Hinreiner E, Webb A. Evaluation of thresholds and minimum difference concentrations for various constituents of wines. 1. Water solutions of pure substances. *Food Technology*. 1955; 9:23–26.
- Breslin PAS, Gilmore MM, Beauchamp GK, Green BG. Psychophysical evidence that oral astringency is a tactile sensation. *Chemical Senses*. 1993; 18:405–417.
- Cadena RS, et al. Comparison of rapid sensory characterization methodologies for the development of functional yogurts. *Food Research International*. 2014
- Cadoret M, Lê S, Pagès J. A Factorial Approach for Sorting Task data *Food Quality and Preference*. 2009; 20:410–417.
- Cartier R, et al. Sorting procedure as an alternative to quantitative descriptive analysis to obtain a product sensory map. *Food Quality and Preference*. 2006; 17:562–571.
- Chambers E, Wolf MB. Sensory testing methods. In: American Society for Testing and Materials. 1996
- Coxon AMP. Sorting data: Collection and analysis. Sage. 1999
- De Leeuw J, Mair P. Multidimensional scaling using majorization: SMACOF in R. Department of Statistics. UCLA. 2011
- de Saldamando L, Delgado J, Herencia P, Giménez A, Ares G. Polarized sensory positioning: Do conclusions depend on the poles? *Food Quality and Preference*. 2013; 29:25–32.
- Dehlholm C, Brockhoff PB, Meinert L, Aaslyng MD, Bredie WL. Rapid descriptive sensory methods-comparison of free multiple sorting, partial napping, napping, flash profiling and conventional profiling. *Food Quality and Preference*. 2012; 26:267–277.
- Dooley L, Lee Y-s, Meullenet J-F. The application of check-all-that-apply (CATA) consumer profiling to preference mapping of vanilla ice cream and its comparison to classical external preference mapping. *Food Quality and Preference*. 2010; 21:394–401.
- Falahee M, MacRae AW. Perceptual variation among drinking waters: The reliability of sorting and ranking data for multidimensional scaling. *Food Quality and Preference*. 1997; 8:389–394.

- Gawel R, Iland PG, Francis IL. Characterizing the astringency of red wine: a case study. *Food Quality and Preference*. 2001; 12:83–94.
- Gawel R, Oberholster A, Francis IL. A ‘Mouth-feel Wheel’: terminology for communicating the mouth-feel characteristics of red wine. *Australian Journal of Grape and Wine Research*. 2000; 6:203–207.
- Green BG. Oral astringency: a tactile component of flavor. *Acta Psychologica*. 1993; 84:119–125. [PubMed: 8237452]
- Joslyn M, Goldstein JL. Astringency of fruits and fruit products in relation to phenolic content. *Advances in Food Research*. 1964; 13:179–217. [PubMed: 14280866]
- Kallithraka S, Bakker J, Clifford MN. Evaluation of bitterness and astringency of (+)-catechin and (–)-epicatechin in red wine and model solution. *Journal of Sensory Studies*. 1997; 12:25–37.
- Kielhorn S, Thorngate JH. Oral sensations associated with the flavan-3-ols (+)-catechin and (–)-epicatechin. *Food Quality and Preference*. 1999; 10:109–116.
- King M, Cliff MA, Hall JW. Comparison of projective mapping and sorting data collection and multivariate methodologies for identification of similarity-of-use of snack bars *Journal of Sensory Studies*. 1998; 13:347–358.
- Lawless, HT. *Quantitative sensory analysis: psychophysics, models and intelligent design*. John Wiley & Sons; 2013.
- Lawless HT, Glatter S, Hohn C. Context-dependent changes in the perception of odor quality. *Chemical Senses*. 1991; 16:349–360.
- Lawless HT, Horne J, Giasi P. Astringency of organic acids is related to pH. *Chemical Senses*. 1996; 21:397–403. [PubMed: 8866103]
- Lawless HT, Sheng N, Knoops SS. Multidimensional scaling of sorting data applied to cheese perception. *Food Quality and Preference*. 1995; 6:91–98.
- Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. *Journal of statistical software*. 2008a; 25:1–18.
- Lê S, Pagès J, Husson F. Methodology for the comparison of sensory profiles provided by several panels: Application to a cross-cultural study. *Food Qual Prefer*. 2008b; 19:179–184.
- Lee CB, Lawless HT. Time-course of astringent sensations. *Chemical Senses*. 1991; 16:225–238.
- Lelièvre M, Chollet S, Abdi H, Valentin D. What is the validity of the sorting task for describing beers? A study using trained and untrained assessors. *Food Quality and Preference*. 2008; 19:697–703.
- Lim J, Lawless HT. Qualitative Differences of Divalent Salts: Multidimensional Scaling and Cluster Analysis. *Chemical Senses*. 2005; 30:719–726. [PubMed: 16221799]
- Meyners M, Castura JC, Carr BT. Existing and new approaches for the analysis of CATA data. *Food Quality and Preference*. 2013; 30:309–319.
- Miller DM, Wiley DE, Wolfe RG. Categorization methodology: An approach to the collection and analysis of certain classes of qualitative information. *Multivariate Behavioral Research*. 1986; 21:135–167.
- Murray J, Delahunty C, Baxter I. Descriptive sensory analysis: past, present and future. *Food research international*. 2001; 34:461–471.
- Murtagh, F. *Multidimensional clustering algorithms*. Vol. 1985. *Compstat Lectures*, Vienna: Physika Verlag; 1985. p. 1
- Nestrud MA, Lawless HT. Perceptual mapping of citrus juices using projective mapping and profiling data from culinary professionals and consumers. *Food Quality and Preference*. 2008; 19:431–438.
- Nestrud MA, Lawless HT. Perceptual mapping of apples and cheeses using projective mapping and sorting. *Journal of Sensory Studies*. 2010; 25:390–405.
- Peleg H, Bodine KK, Noble AC. The influence of acid on astringency of alum and phenolic compounds. *Chemical Senses*. 1998; 23:371–378. [PubMed: 9669051]
- Rasinski KA, MINGAY D, BRADBURN NM. Do respondents really “mark all that apply” on self-administered questions? *Public Opinion Quarterly*. 1994; 58:400–408.

- Reinbach HC, Giacalone D, Ribeiro LM, Bredie WLP, Frøst MB. Comparison of three sensory profiling methods based on consumer perception: CATA, CATA with intensity and Napping®. *Food Quality and Preference*. 2014; 32:160–166. Part B.
- Robert P, Escoufier Y. A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Applied statistics*. 1976:257–265.
- Rodriguez M, Albertengo L, Vitale I, Agullo E. Relationship Between Astringency and Chitosan - Saliva Solutions Turbidity at Different pH. *Journal of Food Science*. 2003; 68:665–667.
- Rubico SM, McDaniel MR. Sensory evaluation of acids by free-choice profiling. *Chemical Senses*. 1992; 17:273–289.
- Sinopoli DA, Lawless HT. Taste Properties of Potassium Chloride Alone and in Mixtures with Sodium Chloride Using a Check - All - That - Apply Method. *Journal of food science*. 2012; 77:S319–S322. [PubMed: 22901084]
- Sowalsky RA, Noble AC. Comparison of the effects of concentration, pH and anion species on astringency and sourness of organic acids. *Chemical senses*. 1998; 23:343–349. [PubMed: 9669047]
- Teillet E, Schlich P, Urbano C, Cordelle S, Guichard E. Sensory methodologies and the taste of water. *Food Quality and Preference*. 2010; 21:967–976.
- Thomas CJC, Lawless HT. Astringent subqualities in acids. *Chemical Senses*. 1995; 20:593–600. [PubMed: 8788093]
- Valentin D, Chollet S, Lelievre M, Abdi H. Quick and dirty but still pretty good: A review of new descriptive methods in food science. *International Journal of Food Science & Technology*. 2012; 47:1563–1578.
- Varela P, Ares G. Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Research International*. 2012; 48:893–908.

Highlights

- Three rapid profiling methods – PSP, CATA and sorting were compared
- Astringent stimuli included multivalent salts, organic acids and polyphenols
- The resulting perceptual maps showed significant similarity across methods
- Groups in the maps appeared to be driven primarily by side tastes of the stimuli

Figure 1.

Visual maps from CATA (A), sorting (B), and PSP (C). Data were analyzed using correspondence analysis (CA), multidimensional scaling (MDS), and multiple factor analysis (MFA) for CATA, sorting, and PSP, respectively. In panel A, the CATA terms endorsed by less than 25% of participants are shown in grey, and the precise location of the terms is indicated by the cross symbol. A composite plot calculated using MFA on the two-dimensional solutions from the three experiments is also included (D). Stimuli abbreviations are the same as in Table 1.

Figure 2. Dendrograms from CATA (A), sorting (B), and PSP (C) experiments. Agglomerative hierarchical clustering was conducted using Ward's minimum variance as the linkage criteria. Stimuli abbreviations are the same as in Table 1. .

Table 1

Astringent compounds used in Experiments 1, 2 & 3. MS= multivalent salt, OA=organic acid, P= polyphenol

Compound	Conc (g/L)	Class	Abbreviation
Ammonium Aluminum Sulfate	2.75	MS	Alum
Zinc Chloride	4.00	MS	ZnCl ₂
Lactic Acid	3.00	OA	Lact-Acid
Malic Acid	2.20	OA	Mal-Acid
Tartaric Acid	1.60	OA	Tar-Acid
Biotan® ¹	5.00	P	Biotan
CocoaVia® ²	32.50	P	Co-Via
Cranberry Extract	1.50	P	Cran
Epigallocatechin Gallate	2.20	P	EGCG
Tannic Acid	2.00	P	Tan-Acid

¹Commercial grape tannin extract used in the wine industry.

²Cocoa extract supplement; Dark Chocolate Unsweetened flavor.

Table 2

Terms included in the check-all-that apply (CATA) question. Presentation order of terms was counter-balanced in software using a Williams Design. To avoid dumping, 'Other' was included, along with an open-ended text box.

Check all attributes which describe the sample:
☐ Bitter
☐ Burning
☐ Drying
☐ 'Other'
☐ Puckering
☐ Roughing
☐ Salty
☐ Sour
☐ Stinging/Pricking
☐ Sweet
☐ Tickle
☐ Umami/Savory
☐ Velvety

Table 3

NRV coefficients and p-values between sample configurations in the first two dimensions of multivariate statistical techniques for the three experiments. NRV coefficient is interpreted in the same way as a z-score for a normal distribution, with a large score (>2) indicating significant similarity between configurations. As would be expected, the composite MFA from the 3 different methods was significantly similar with the 3 underlying 2D configurations on which it was based (all NRV's > 5.5 ; p's < 0.0005).

	CATA	Sorting	PSP
CATA	1	5.71(0.0005)	6.07(0.0003)
Sorting	-	1	5.81(0.0005)
PSP	-	-	1

Table 4

Percentage of participants who selected each of the terms of the check-all-that-apply (CATA) experiment. Cochran's Q Test was used to determine level of significance across samples with respect to terms.

CATA Term	Compound											
	Multivalent Salts				Organic Acids				Polyphenols			
	Alum1	Alum2	ZnCl2	Lact-Acid	Mal-Acid	Tart-Acid	Biotan	Co-Via	Cran.	EGCG	Tan-Acid	
Bitter**	46.3	39.0	46.3	26.8	26.8	22.0	68.9	51.2	82.2	75.6	85.4	
Burning*	0	0	12.2	4.90	2.40	2.40	2.40	0	2.40	17.1	12.2	
Drying**	70.7	75.6	73.2	39.0	41.5	39.0	75.6	41.5	51.2	48.8	61.0	
“Other”*	4.90	4.90	12.2	2.40	2.40	2.40	12.2	22.0	9.80	14.6	4.90	
Puckering**	34.1	26.8	19.5	51.2	51.2	46.3	12.2	7.30	12.2	19.5	12.2	
Roughing*	29.3	31.7	29.3	14.6	14.6	12.2	39.0	26.8	24.4	29.3	29.3	
Salty**	4.90	7.30	39.0	26.8	7.30	14.6	7.30	4.30	9.80	0	4.90	
Sour***	43.9	48.8	14.6	87.8	80.5	80.5	19.5	12.2	17.1	9.80	14.6	
Stinging/Pricking*	4.90	4.90	17.1	12.2	7.30	7.30	7.30	0	12.2	22.0	12.2	
Sweet***	29.3	29.3	12.2	22.0	14.6	14.6	2.40	4.90	2.40	0	0	
Tickle	2.40	2.40	7.30	7.30	9.80	4.90	0	7.30	0	2.40	0	
Unami/Savory	2.40	7.30	12.2	9.80	2.40	0	7.30	9.80	2.40	2.40	2.40	
Velvety***	7.30	12.2	12.2	0	2.40	9.80	12.2	31.7	2.40	2.40	4.90	

* Indicates significant differences at $\alpha=0.05$;

** $\alpha=0.01$;

*** $\alpha=0.001$ across samples using Cochran's Q test