COMPARISON OF RESULTS FROM CUPPING AND DESCRIPTIVE SENSORY ANALYSIS OF COLOMBIAN BREWED COFFEE

B. DI DONFRANCESCO¹, N. GUTIERREZ GUZMAN² and E. CHAMBERS IV¹,³

¹The Sensory Analysis Center, Kansas State University, Manhattan, KS 66502
²Facultad de Ingeniería, Universidad Surcolombiana, Neiva, Huila, Colombia
³Corresponding author.
TEL: +1-785-532-0156;
EMAIL: eciv@ksu.edu
Accepted for Publication June 24, 2014
doi:10.1111/joss.12104

ABSTRACT

Sensory profiles of 13 coffee samples from the Huila Region, Colombia were evaluated using two different sensory panels: a highly trained descriptive sensory panel and a group of Q-certified coffee cuppers. The trained panel consisted of six descriptive panelists who developed a lexicon to evaluate and then test the coffee samples. Four “cuppers” scored the same samples based on the Specialty Coffee Association of America “cupping protocol.” In addition, cuppers generated tasting notes to characterize the different coffee samples. Data indicated little overlap between the two methods and a low relationship between the two different sets of terms. Moreover, tasting notes by cuppers indicate lack of agreement on the terms used to describe samples, with only four terms used by more than two assessors to describe a single coffee product out of a total of 59 terms used by the cuppers. The results indicate that the cupping method and sensory descriptive methods provide different information that cannot be used as an alternative to each other when describing coffee products. Instead, the results suggest that the two types of data may be used synergistically to evaluate the quality and the sensory properties to better characterize coffee samples.

PRACTICAL APPLICATIONS

This research shows that “expert” coffee cupper data and trained sensory panel data cannot be used interchangeably. Thus, for research purposes sensory panel information is necessary for tracking changes in sensory properties.

INTRODUCTION

Coffee, after water, is the most popular beverage on earth. Factors such as the plant variety, the geographical areas where coffee is grown, climatic conditions, processing methods before the roasting phase, roasting speed and level, grind size and brewing methods, participate in creating the different sensory properties existing from one product to another (Bhumiratana et al. 2011). The range of geographically different markets where this product is traded can be wide and the techniques through which coffee quality is defined by international and local institutions as well as trading and private companies vary tremendously (Feria-Morales 2002).

Trained panelists for descriptive analysis and expert tasters for flavor evaluation and quality control both are commonly used by industries. Training enhances the ability to recognize, describe and measure the intensity of volatiles by individuals with reasonable sensory acuity and the ability to focus their attention on specific sensory stimuli (Lawless 1984).

Sensory methods, rather than instrumental methods, are most frequently used to estimate aroma, taste and flavor of coffee brews (Nebesny and Budryn 2006; Alvarado and Linneman 2010). In coffee beverages, as in many products, the presence of several hundred volatile compounds can make pinpointing specific compounds responsible for a flavor character difficult (Sanz et al. 2002; Chambers and Koppel 2013).

Industries such as coffee, perfume, tea and tobacco often use “experts” during the quality assessment process. This type of assessor can be defined as persons who accumulate, year after year, wide knowledge about a specific product, and are involved in the decision-making process by
companies (Gatchalian 1981). Lawless (1984) categorized experts as assessors who are not trained panelists, but have long-standing experience with a specific product and are used more for quality control or to develop new products such as perfumers and flavor chemists. Professional coffee judges, often called cuppers, are widely used around the world. These experts often are sensitive to any change in the characteristics of the product (Feria-Morales 2002).

However, the use of experts presents some problems. Bias from knowing the particular sample, the influence of external factors, change in perceptive abilities of an individual through illness, etc., and the long time that can be necessary to develop this kind of professional expertise discourage their use as the sole evaluation tool. Using a panel of trained sensory assessors for a detailed evaluation of raw materials or products following a rapid initial assessment by the experts is increasing (Feria-Morales 2002). Trained descriptive panels are widely used in many industries as the large volume of literature in this area attests.

Sensory lexicons are used by trained descriptive panelists to help standardize the characteristics of a product across different people/panels (Lawless and Civille 2013). Thus, lexicons are used for descriptive sensory panels, experts and consumers. Different characteristics are used in the sensory vocabulary, depending on the level of training and experience of the assessors.

Conventional sensory analysis of coffee using well-trained panelists tries to deconstruct flavor in coffee (e.g., Narain et al. 2003) and may use a number of different characteristics. For example, for some complex products such as soy sauce (Cherdchu et al. 2013) or dog food (Di Donfrancesco et al. 2012) or broad ranges of products such as spices (Lawless et al. 2012), sensory panels may develop 50–100 descriptors although these may not be used in every evaluation. Products that are less complex or have a more limited range of product options such as turnip greens (Arias-Carmona et al. 2012) and kimchi (Chambers et al. 2012) typically may have 20 or fewer attributes of interest.

Expert coffee tasters like “cuppers”, who are trained to distinguish among small differences of coffee beverages, use a reasonably large vocabulary as can be seen in the various coffee wheels that have been developed (e.g., SCAA 1995 or Counter Culture Coffee 2013). Those wheels show a large number of coffee attributes (>50) that can be used in cupping.

For coffee, Hayakawa et al. (2010) produced a large list of more than 100 descriptors for appearance, aroma, taste/flavor, mouthfeel and overall impression using cuppers and experienced coffee professionals. Trained panelists also evaluated the samples, but did not create any additional attributes to those from the experts. Consumers in that study used far fewer terms to describe coffee, with only 29 terms used by more than 50% of the consumers to describe coffee. Jervis et al. (2012) also used a small number of sensory attributes with consumers in a study of latte-style coffees.

The most commonly used method now as well as in the past to evaluate coffee quality in the cup is called “cupping”. The grading system developed by the Specialty Coffee Association of America (SCAA 2009) includes a list of standard attributes, such as body and acidity, to be scored in order to describe the product. Cuppers also are encouraged to add “tasting notes” that provide additional information on the specific characteristics of the coffee. However, one of the problems in coffee evaluation by expert “cuppers” is the lack of a common language for flavor notes and a consistent vocabulary across cuppers. Tasting notes can help to better describe and compare flavor notes essential in discriminating among different coffee beverages beyond the main defects (Castle 1986).

Vocabularies (ICO 2010) and handbooks (Castle 1986; Lingle 2001) have been developed to list terms that can describe sensory properties of coffee. Often these vocabularies focus on trying to develop terms influenced by specific cultural and linguistic aspects in order to be also recognized by local consumers of a given country (Seo et al. 2009; Hayakawa et al. 2010). Although helpful to marketing, such vocabularies are difficult to translate across cultures and include many terms that are overly broad, poorly defined or not well understood by others, which makes this a problem for research.

The objective of this study was to compare the similarity of information found for a selection of Colombian coffees as tested by a trained sensory panel and “cuppers” working for the coffee industry in Colombia.

MATERIALS AND METHODS

Samples

Thirteen Arabica coffee bean samples from Colombia, specifically from the Pitalito area in the Huila Region, were used in the study. Within the Pitalito area, samples were from different subareas. Four samples originated from the Southern area, five samples from the Middle area and four samples from the Northern area. They came from various altitudes. All samples were prepared to a medium roast. Samples used for the cupping method were evaluated within the coffee factory where they were processed, in a room dedicated to the evaluation. Products for the sensory test were obtained about 3 weeks before testing and were roasted at the same time as those used for cupping. Samples were stored in the Sensory Analysis Center at Kansas State University in sealed polyvinyl chloride (PVC) nonpermeable bags further contained in plastic boxes at room temperature.
Sample Preparation

For both tests, coffee beans were ground no more than 30 min before infusion. The amount of coffee used for the infusion preparation during the evaluation was established at 5.5 g per 100 mL of water. This ratio is in accord with the International Standard for the preparation of coffee samples for use in sensory analysis (ISO 6668 2008) that suggests a ratio in a 5–9 g range per 100 mL of water and was the same ratio the Colombian cuppers used. Evaluation was conducted in a quiet environment, with no interfering aromas in each of the locations.

Descriptive Sensory Analysis. For the descriptive sensory studies, coffee was ground to a “medium” grind and hot (−98°C) purified (reverse osmosis, carbon filtered) water was poured over the ground coffee. The infusion was then allowed to steep for 3 min before being filtered through a fine mesh, metallic tea strainer. Only the liquid brew, without grounds, was presented to the panelists.

Cupping Method. For the cupping method, the tasting protocol followed the SCAA protocol (SCAA 2009). Coffee samples were roasted within 24 h of evaluation and allowed to rest for at least 8 h before grinding just before the evaluation. For sample uniformity evaluation, five cups of each sample were prepared. Clean and odor-free water was used, with ideal dissolved solids in the 125–175 ppm range. Water was brought to approximately 92–94°C and poured directly on the ground coffee. Grounds were steeped for 3–5 min before the sample was evaluated. The entire cup was presented to the cuppers, including the grounds, which formed a “crust” on top. The procedure calls for smelling, breaking the crust, re-smelling and tasting the coffee when the grounds have mostly sunk to the bottom of the cup.

Panelists

Descriptive Analysis. Six highly trained panelists from the Sensory Analysis Center, Kansas State University (Manhattan, KS) took part in this study. All of the panelists had completed 120 h of general descriptive analysis panel training with a wide range of different food products. Techniques and practice in attribute identification, terminology development and intensity scoring were part of the training. Panelists participating in the study had more than 1,000 h of testing experience with a variety of food products. Further orientation focusing on coffee was provided to panelists, using samples included in the study as well as other different samples.

Cupping Method. The four Q-certified cuppers evaluating samples in this study were official cuppers with several years of experience in the Colombian coffee industry.

Terminology

Descriptive Sensory Analysis. Panels were asked to develop a lexicon based on the attributes present in the sample set in the study. The terminology was developed for aroma, flavor and aftertaste. During an initial phase, several coffee samples, including samples not part of the study, were shown to the panelists. This step was considered necessary in order to conduct a preliminary orientation to the product category. All the samples that were the object of study were shown to the panel during the terminology development stage. Various references were provided to the panelists; some references were initially proposed by the panel members and researchers according to previous experience and prior studies (e.g., Hayakawa et al. 2010). Other references were added during the terminology development phase.

More than 10 1.5-h sessions were held to establish, discuss and refine the attributes definitions to be included in the final terminology and the descriptive references. The goal of this phase was to try to avoid redundant attributes and gross overlap among descriptive terms. The panel had this list of terms, definitions and references during the entire evaluation process and can refer to it at any time.

Cupping Method. The cupping form used includes terms considered important flavor attributes for coffee. Terms such as fragrance/aroma, flavor, aftertaste, acidity, body, balance, uniformity, clean cup and sweetness are considered positive scores of quality according to a rate by the cuppers. The attribute defects reflects a negative score, indicating unpleasant flavor sensations. The term overall is scored according to the flavor experience of individual cuppers using their own personal appraisal. After scoring each of those terms, the cuppers include specific individual attributes that they felt described the flavor of the coffee. Usually from one to five “descriptive” terms were used by the cuppers at this point.

Sample Evaluation Procedure

Descriptive Sensory Analysis. Ten 2-h sessions were held for the evaluation stage that included both aroma and flavor attributes. Three to four samples were evaluated during each session. A three-digit random code identified each sample and the order in which the samples were evaluated was randomized. To measure the intensity, a numerical scale of 0–15 with 0.5 increments, where 0 represented none and 15 extremely high, was applied. Each sample was
individually evaluated by each panelist on a ballot based on aroma and flavor references listed in the terminology previously developed. Three replicates were held for each of the 13 samples. The coffee sample was transferred to an insulated hot pot and brought into the panel room during the evaluation. The serving temperature was set to a range of 60–65°C, based on input from the panelists during orientation for the temperature level that worked best for both aroma and flavor evaluation. Aroma evaluation was conducted first for each sample. The infusion was poured into 120 mL (4 oz) Styrofoam cups (Dart, Mason, MI, USA) and was filled to approximately one-half (60 mL) of the total volume. The headspace was needed for aroma evaluation based on panelist comments. When the panelists were ready for flavor evaluation, a second cup was filled to approximately three-fourths (90 mL) of the cup in order to have an adequate amount of sample to be tasted.

**Cupping Method.** According to the SCAA tasting protocol, the scale used is a 16-point scale representing levels of quality from 6 to 9.75, with quarter point increments. A range from 6.00 to 6.75 is considered as Good, from 7.00 to 7.75 as Very Good, from 8.00 to 8.75 as Excellent, and from 9.00 to 9.75 (10.00 in this study) is considered Outstanding (SCAA 2009). In theory, the scale ranges from a minimum of 0 to a maximum of 10, but scores below 6 or a full 10 are rarely, if ever, given. In practice, the cuppers in this study did not score below 6, but did score the full 10 points when they considered this necessary.

Attributes were rated following a specific order, according to the flavor perception changes caused by the decreasing temperature of the sample. **Fragrance/aroma** was the first attribute to be evaluated. First, within 15 min after the coffee beans were ground, the fragrance was evaluated. Later, 3 min after infusing with water but no more than 5 min, the aroma is evaluated again. Then, the surface crust was broken and cuppers evaluated the aroma a third time. Although cuppers evaluate these three steps separately, at the end a unique **fragrance/aroma** score on the sheet is required, reflecting both dry and wet evaluation. After about 10 min from infusion, when the sample cools down to a temperature close to 70°C (160°F), **flavor** and **aftertaste** were evaluated. These attributes are evaluated at this time because at this temperature the retronasal vapors are assumed to be at their maximum intensity. The procedure cuppers follow to do this consists of aspirating the sample into the mouth, slurping in a way that permits the liquid to cover as much area as possible, particularly the tongue and the upper palate. The next attributes evaluated when the temperature continues to cool down (70–60°C) were **acidity**, **body** and **balance**. The last attribute reflects how **flavor**, **aftertaste**, **acidity** and **body** work together. Assessors repeat the evaluation of all the attributes above for a further 2–3 times while the temperature continues to decrease. When temperature reaches levels below 37°C, **sweetness**, **uniformity** and **clean cup** are evaluated. For these attributes each of the five cups of the same sample is individually scored, earning a maximum of two points per cup. When the temperature reaches a temperature close to 20°C (70°F) the evaluation ends. At this point the cupper gives an **overall** score based on all the combined previously scored attributes. The final step consists of adding all the scores and detracting the eventual negative or poor flavors indicated as **defects**. This scoring procedure follows particular rules described in the SCAA protocols. Defects are classified in taints and faults. The first is an off-flavor that is noticeable but not overwhelming and this is given a “2” score in intensity. Faults are overwhelming off-flavors that can make the sample unpalatable. This is given a “4” as an intensity score.

The final scores are then given from the total scores minus the defects scores. The final score has a maximum of 100, and the score ranges have been classified as 90–100, indicating an outstanding product, 85–89.99 an excellent product, and 80–84.99 considered a very good product. Coffee samples belonging to these three categories are indicated as “speciality.” Samples earning a total score below 80 are not considered “speciality” coffee.

**Data Analysis**

Data from all procedures were averaged over panelists/cuppers and replications (descriptive analysis) and used in a principal components analysis (PCA), conducted in Unscrambler version 10.2 (Camo Software AS, Oslo, Norway). For descriptive analysis, a PCA was conducted for aroma and separately for flavor and aftertaste attributes scores because of the number of attributes and the repetition of attributes in aroma and flavor. For the cupping scores, aroma, flavor and aftertaste are combined. For the descriptive analysis, attributes receiving a score <1.0 for all products and that did not differ significantly among products were not included in the PCA. For aroma, the attributes malt, spice brown, pepper, pungent, brown sweet, caramelized, vanilla, honey, molasses, chocolate, floral, fruity, fruity dark, fruity citrus, fruity noncitrus, fruity berry, overripe near fermented, green, green pea pod, green herb-like, fermented, musty earthy, moldy damp, stale, rubber-like and medicinal earned scores <1.0 for all of the samples (data not shown). For flavor, the attributes earning a score <1.0 were malt, spice brown, pepper, brown sweet, caramelized, vanilla, honey, syrup, molasses, chocolate, floral, fruity dark, fruity citrus, fruity berry, fruity noncitrus, overripe near fermented, raw, green pea pod, green herb-like, green hay-like, fermented, musty earthy, moldy damp, stale and sweet.
These attributes were not included in the PCA analysis except for musty earthy and metallic that showed, for flavor, a significant difference (P-value < 0.05) among the different samples. To have consistency, musty earthy was kept in the PCA analysis for aroma too.

Unscrambler also was used for partial least square regression (PLSR) and correlation analysis to compare cupping and descriptive analysis methods.

In addition to those analyses, for cupping, the number of times a particular flavor attribute was used was counted for each coffee. Thus, a maximum number of “4” would indicate that all cuppers used the same term for that coffee. Alternatively, a minimum number of “1” would indicate that note was used by only one cupper.

RESULTS AND DISCUSSION

Terminology

Descriptive Terminology Development. At the end of the terminology development sessions, a total of 92 terms were agreed on by all panelists. Forty-three attributes were related to aroma, 46 to flavor and three to aftertaste. Additional aftertaste attributes existed (all included in flavor) but it was decided to use only the narrow group of three aftertaste attributes that the panelists generally detected in all the samples in the study. For each attribute, definitions and references were generated. All the attributes are listed in Table 1.

Some aroma attributes such as pepper, honey, molasses, fruity citrus, fruity berry, green herb-like and fermented were detected only in three or fewer samples of the 13. For flavor, pepper, honey and syrup were detected in three or fewer samples.

Cupping Terminology – Tasting Notes. The terms used in the tasting notes by cuppers are shown in Table 2. There is high inconsistency among cuppers’ notes. Of a total of 59 tasting notes, only four were used by more than two cuppers for a specific sample. This is problematic if tasting notes are used to provide information to processors or consumers, as often is done on website or packaging. The lack of consistency in tasting notes indicates that different cuppers would use different terms to describe the key aroma/flavor notes in the coffee, leading to confusion as to what the coffee actually tasted like. This inconsistency probably can be explained by a lack of discussion and agreement on common definitions related to the terms.

Coffee Comparisons

Descriptive Analysis. For aroma, Factors 1–4 explained 80.7% of the variability with Factor 1 explaining 36.2% of the variability and Factor 2 explaining 25.9%. The PCA plots related to aroma attributes for Factors 1 and 2 are shown in Fig. 1. Factor 1 was positively related to acrid, ashy and burnt aroma, and negatively related to grain, sweet aromatics and nutty aroma attributes. Factor 2 was positively related to grain, cocoa and sweet aromatics, and negatively related to sour aromatics and cardboard. Factor 3, which explained 10.8% of the variability (not shown), was positively related to the musty/earthy note. Factor 4 explained 7.8% of variability (not shown) and was positively related to
**TABLE 2.** TASTING NOTES GENERATED BY CUPPERS TO CHARACTERIZE THE SAMPLES. NOTES ARE DIVIDED ACCORDING TO THE MAXIMUM NUMBER OF ASSESSORS USING THE NOTE TO DESCRIBE A SINGLE SAMPLE (ON A TOTAL OF FOUR EVALUATING CUPPERS)

<table>
<thead>
<tr>
<th>Used by three cuppers</th>
<th>Used by two cuppers</th>
<th>Used by one cupper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citric</td>
<td>Fruit</td>
<td>Sweet fruit</td>
</tr>
<tr>
<td>Sweet</td>
<td>Yellow fruits (peach, mango, papaya)</td>
<td>Mango</td>
</tr>
<tr>
<td>Candy</td>
<td>Red fruits (strawberry, cherry)</td>
<td>Fresh lemon</td>
</tr>
<tr>
<td>Malt</td>
<td>Honey</td>
<td>Lime</td>
</tr>
<tr>
<td></td>
<td>Sugarcane</td>
<td>Fresh fruits</td>
</tr>
<tr>
<td></td>
<td>Grass</td>
<td>Raisin grape</td>
</tr>
<tr>
<td></td>
<td>Rough</td>
<td>Mandarin</td>
</tr>
<tr>
<td></td>
<td>Chocolate</td>
<td>Mild coffee</td>
</tr>
<tr>
<td></td>
<td>Nut</td>
<td>Bitter chocolate</td>
</tr>
<tr>
<td></td>
<td>Hazelnut</td>
<td>Cocoa</td>
</tr>
<tr>
<td></td>
<td>Flower</td>
<td>Creamy</td>
</tr>
<tr>
<td></td>
<td>Astringent</td>
<td>Creamy body</td>
</tr>
<tr>
<td></td>
<td>Strong body</td>
<td>Roast nut</td>
</tr>
<tr>
<td></td>
<td>Balance</td>
<td>Peanut</td>
</tr>
</tbody>
</table>

**FIG. 1.** PRINCIPAL COMPONENTS ANALYSIS BILOT OF DESCRIPTIVE SENSORY AROMA FOR FACTOR 1 (F1) AND FACTOR 2 (F2) SHOWING PRODUCTS (BOXES) AND ATTRIBUTE ABBREVIATIONS
cardboard aroma, and negatively with nutty, roasted and burnt aroma notes. It should be noted in the aroma PCA plots that differentiation of the samples is low in this set of data, ranging from −0.65 to +0.8 on the PCA factor coordinates. This suggests that although these samples can be differentiated for aroma, the samples still have reasonable similarity.

For flavor and aftertaste, Factors 1–4 explained 87.7% of the variability with Factor 1 explaining 49.6% of the variability and Factor 2 explaining 23.3%. PCA plots of those two factors for flavor are shown in Fig. 2. Factor 1 was positively related to several attributes such as acrid, burnt, ashy, bitter flavor and bitter aftertaste attributes, and negatively with sweet aromatics. Factor 2 was positively related to bitter aftertaste, and negatively related to smoky and ashy flavor notes. Factor 3 explained 8.2% of the variability (not shown) and was positively related to bitter. Factor 4, which explained 6.7% of the variability (not shown), was positively related to musty/earthy, and negatively with musty/dusty. The PCA coordinates for the samples based on flavor are larger, indicating that the differentiation of flavor is greater than for aroma.

**Cupping.** Principal component analysis for cupping evaluation resulted in Factors 1–4 explaining 89.9% of variation with Factor 1 explaining 42.6% of variation and
Factor 2 explaining 24.0% (Fig. 3). Factor 1 was mostly related to balance, but it must be noted that all of the attributes are grouped in an area positively related to Factor 1, except for uniformity. This suggests that majority of the attributes scored by cuppers overlap and, at least in this study, did not provide much information beyond that given by balance and uniformity. Factor 2 was positively related to uniformity. Factor 3, explaining 15.8% of the variability, was positively related to balance and negatively related to acidity (not shown). Factor 4, explaining 7.5% of the variability, was positively related to attributes such as clean cup and fragrance/aroma.

The cuppers clearly were able to differentiate the samples using these quality characteristics with several samples seeming to score quite low on Factor 1 (balance) and other samples scoring considerably higher. Similarly, several products scored low for Factor 2 (uniformity) while others scored much higher.

Relationships Among Cupping Data and Descriptive Sensory Panel Data

In the PLSR, cupping method scores (x) were compared with descriptive analysis scores (y). The related plot is shown in Fig. 4. The plot clearly demonstrates that with only a little exception, the two sets of evaluations (cuppers and trained descriptive panelists) are quiet distinct and do not show any particular relationship. Some overlapping can be observed between the majority of the cupping attributes and some descriptive aroma attributes such as cocoa, grain,
nutty and sweet aromatics as well as the flavor attributes musty dusty, woody and roasted. Uniformity from the cuppers appears in a similar part of the map to sweet aromatics scored by the descriptive panel.

Correlation analysis indicated only a few moderate correlations. There was a positive correlation ($r = 0.66$) between grain aroma (descriptive) and fragrance/aroma (cupping), and negative correlations between sour aromatics (descriptive) and fragrance/aroma (cupping) ($r = -0.77$), and musty/earthy flavor (descriptive) and aftertaste (cupping) ($r = -0.66$).

Although mathematically there appear to be some minimal relationships based on the PLSR and correlation, realistically those seem to be anomalies of the sample that likely have little actual relationship. It is unlikely that balance scores from cuppers depend on sensory properties such as cocoa, grain, nutty, sweet aromatics, and musty dusty, woody and roasted flavors. Similarly, uniformity among cups (cupping) and sweet aromatics (descriptive) likely are not actually highly related. Further studies with a larger range of samples would need to be conducted to determine whether there actually are stable relationships among those characteristics measured by the two groups.

What is clear from looking at the data is that the cupping attributes generally are far from the descriptive sensory attributes. This indicates little relationship among the individual characteristics measured by a trained sensory panel and the more broad quality characteristics measured by cuppers. Thus one panel cannot replace the other regardless of the task. Neither can one type of data be used at this point to predict the other. Much more research would need to be done to establish a relationship and to determine if either of these methods relates to consumer acceptance. This data suggest that the panels may need to work synergistically with each other to provide the broadest understanding of the coffee samples.

Another way to examine the data is to look at the differentiation among the samples by the two groups. Using Figs. 2 and 3 it is apparent that both groups differentiated the samples, but they did so in different ways. For example, samples ZM322 and ZM327 show as quite different from other samples in the descriptive panel data (Fig. 2). However, using cupping data (Fig. 3) ZN322 is seen as more similar to ZN332, ZS319 and ZM330 than with ZM327. Similarly, cupping data (Fig. 3) suggested that ZS335 and ZM334 were quite different from each other and different from other samples, whereas in the descriptive panel data (Fig. 2) ZM335 is more similar to ZN329 and ZM334 is quite similar to ZN326 for flavor. It should be noted that the aroma data (Fig. 1) also show large differences versus cupping, and some generally smaller differences versus the flavor/aftertaste data. Thus using aroma, which might be
considered an option for rapid testing because it typically provides less carryover from sample to sample, likely is inadequate to represent the flavor or the overall quality of the coffee.

CONCLUSIONS

Thirteen coffee samples from beans cultivated in the Huila Region, Colombia were evaluated both by a team of expert “cuppers,” widely used in the coffee industry for quality assessment, and by a trained sensory panel to compare information. A list of terms was generated by sensory panellists to describe the samples for descriptive testing. The SCAA cupping protocol attributes were scored for the cupping method and “tasting notes” of key flavors were collected. Little relationship exists among the two procedures. This is not surprising considering that cupping is designed to evaluate “quality” and descriptive sensory analysis is designed to evaluate specific sensory characteristics. However, this poor relationship suggests that neither method is appropriate for using in all situations. Both methods may need to be used for the greatest understanding of coffee samples under evaluation. Individual cuppers use different terms to describe the same coffee, which provides further evidence that the two methods are dissimilar and can be synergistic, with cupping providing overall measures of quality and descriptive sensory analysis providing the detailed, consistent description that cupping lacks. This study provides a preliminary step for a future work leading to better integration between the “experts” evaluation and descriptive sensory analysis related to coffee.

ACKNOWLEDGMENT

The authors thank Mr. Javier Murgecito, Mild Coffee Company, for providing samples and the Colombian staff essential to the execution of this study.

REFERENCES
