2 Classical Descriptive Analysis

Hildegarde Heymann, Ellena S. King, and Helene Hopfer

CONTENTS

2.1 INTRODUCTION AND HISTORY OF DESCRIPTIVE ANALYSIS

Classical or generic descriptive analysis (DA) is the gold standard technique in sensory science (Lawless and Heymann 2010). The method allows the experimenter to describe all the sensory attributes associated with a product and sensory differences among products. The technique is used extensively, particularly in the food, beverage, and personal care industries, as can be seen from a few examples published in 2012 on food products (Alasalvar et al. 2012; Cakir et al. 2012; Elmaci and Onogur 2012; Paulsen et al. 2012;

Zeppa et al. 2012), beverages (Garcia-Carpintero et al. 2012; Keenan et al. 2012; Ng et al. 2012; Parker et al. 2012; Sokolowsky and Fischer 2012), and other consumer products (Bacci et al. 2012; Verriele et al. 2012).

The current technique of DA originates from three different methods: Flavor Profile (FP®), Quantitative Descriptive Analysis (QDA®), and the Spectrum Method®. FP® was invented by Jean Caul and coworkers (Cairncross and Sjostrom 1950; Sjostrom et al. 1957) when they evaluated the effect of monosodium glutamate on food flavor. In this method, a group of panelists and the panel leader describe the products by consensus using agreed upon terminology and a nonnumerical scale. In the early 1970s, Herbert Stone, Joel Sidel, and others (Stone et al. 1974) created QDA®, which changed the FP® by removing the consensus evaluation of the products and adding a line scale used by each panelist individually, in replicate. This method retained the consensus generation of the attributes but allowed the use of statistical analysis on the data obtained. In the late 1970s, Gail Civille and others (Munoz and Civille 1998) created the Spectrum Method®, which uses absolute scales and attribute lexicons rather than consensus term generation.

In this chapter, we describe the two generic DA techniques that sprang from these predecessors: consensus-trained DA and ballot-trained DA. As will become clear later in this chapter, the major difference between these techniques is in the generation of attributes that the DA panel uses to score perceived intensities of the products. Despite the underlying differences in the training process, it has been shown that the data from different DA panels are very consistent (e.g., Heymann 1994; Lotong et al. 2002; Martin et al. 2000).

2.2 PROCESS OF DESCRIPTIVE ANALYSIS

2.2.1 Experimental Design

As with all studies, the experiment must be designed before the DA can be performed. Since this is not a chapter on experimental design, the following books and chapters would provide an excellent foundation into the design of DA experiments (Gacula et al. 2008; Meullenet et al. 2007, Naes et al. 2010). However, a few key points should always be kept in mind. These are replication, number of panelists, carryover, and number of samples per session. These will be discussed here.

There are scientists who believe that with an extremely well-trained panel, there is no need for replication (Mammasse et al. 2011). However, unless one has spent a great deal of time determining that the panel is truly reproducible (which could take years), it is much better statistically and much faster to add replication in the experimental design. While the

original QDA® suggests four to six replications, general agreement among sensory community is that three replications are sufficient and give enough statistical power, if paired with a trained panel of adequate size.

The literature states that the adequate number of panelists is between 8 and 12 (Lawless and Heymann 2010). A recent publication has also affirmed that this seems to be an ideal number (Heymann et al. 2012). However, the number of panelists may be lower when there are large differences among the samples (Mammasse and Schlich 2012), and conversely, if only subtle differences exist among samples, then more panelists would be required.

If samples are likely to cause a carryover effect from one sample to the next, for example, *astringency* in wines or *heat* in products flavored with chilies, then both an adequate rinsing regime and an experimental design that allow the researcher to determine carryover effects are needed. Using a Williams Latin square design or an incomplete block design for carryover effects for the product presentation in the DA is an easy way to evaluate the effects of sample carryover as needed (Ball 1997; Wakeling and MacFie 1995).

The number of samples served within a session is determined by the type of sample and the specific attributes being evaluated. For samples that are evaluated visually or tactilely, but not orally or nasally, panelist fatigue is less likely, and thus, evaluating up to 15 or 20 samples per session is possible. However, if the samples are evaluated for aroma and flavor, then the number of samples per session should be much lower. If the samples are *challenging*, for example, highly astringent wines, spirits, or very spicy salsas, then the number of samples per session would be even fewer. For example, in a study by Cliff and Heymann (1992), panelists evaluated only three samples in a session during an examination of oral pungency. As a general rule of thumb, about six samples per session seems acceptable.

2.2.2 Panelist Selection

First and foremost, the panelists must be motivated and interested in serving on the panel. If this is true, then we have found in over 30 years of training panelists that essentially everyone can be trained and can be a reliable panelist. Secondly, the panelists must be reliable, in that they arrive when they are supposed to and that they follow instructions. Beyond these requirements, we have found that panelist selection is relatively simple. We usually do not do extensive screening, although others encourage this (e.g., Barcenas et al. 2000; Noronha et al. 1995). However, Nachtsheim et al. (2012) found that screening seems to decrease panel performance—this makes some sense, especially if the screening process is onerous and protracted. The panelists may lose interest and motivation before they even start the training for the specific study. On the other hand, screening for

competence in a specific task may be important for the outcomes of the study. For instance, one should screen for color blindness, if evaluation of color is part of the DA.

In our laboratory, we have also found that trying to train experts, such as expert wine judges, coffee tasters, or dairy judges, as DA panelists can be very frustrating for both the panelist and the panel leader. It is often easier to train panelists who are novices, as far as the specific product is concerned.

2.2.3 Term Generation and Reference Standards

The next step after panel selection is term generation. The procedures differ depending on whether the process will involve consensus training or ballot training.

2.2.3.1 Consensus Training

In this process, the panelists are charged with determining, through consensus, the attributes that discriminate among the samples. On the first day, we usually serve them two or three samples from the product set—these samples are chosen to be as different as possible in order for the panelists to feel that they are actually able to do the task at hand. We then ask the panelists, individually and quietly, without conversation, to determine a list of attributes that discriminate among the given samples. They are told that the attribute terms must be actionable in the sense that we can make a reference standard for it. This means that a term such as *green vegetable* would be acceptable but *yummy* would not. In the case of the last descriptor, it must be made clear to the panel that their opinion or preference for the product is not important. Also ambiguous terms like *complexity* should be discouraged since creating reference standards for such a term would likely be impossible.

Once all panelists have assessed the products, we ask each panelist to read the attributes they used. We write all words on a board—grouping words where possible and indicating words that were used multiple times. This process usually takes about 50–60 min.

At the next training session, we give the panelists another subset of samples from the product set (these are frequently more similar to one another than the first subset) and we repeat the process. During this session, we also start showing the panel potential reference standards to anchor the attributes (see Section 2.2.3.3). The process is repeated as many times as is necessary to allow the panel to see all samples in the product set and to ensure that all potential attribute terms have been listed.

Usually, starting in the third training session, the panel leader will work with the panel to determine which of the listed attributes will

actually be used in the study. There are usually a number of terms that were used by most, if not all, panelists and these clearly need to be in the final attribute list. There are also frequently a number of terms that are synonymous, and in these cases, it is relatively easy to find a compromise. The more problematic terms are the ones that were used by a few panelists but that do seem to describe specific differences among the samples in the product set. For these terms, the panel leader's negotiation skills become crucial. The trick is to minimize the eventual attribute list, to prevent panelist fatigue, while still covering all the differences among the samples. It is often worth adding one or two additional terms to maintain panel harmony, but care must be taken not to add too many. Frequently, an especially vocal proponent of a specific attribute will be mollified if the panel leader explains that the score sheet will have a line scale labeled "Other" where the panelists can indicate the attribute and then score its intensity. The "Other" attribute is also useful to minimize *dumping*. This occurs when panelists perceive a difference in an attribute but the attribute is not part of the listed terms (Lawless and Heymann 2010).

Once the attribute list has been completed, then the training sessions involve making sure that the entire panel is comfortable with the specific reference standards and, most importantly, that they can identify these standards blind. Once all reference standards have been approved and all panelists can identify all standards blind, then the panelists are shown how to use the computerized data acquisition system (if used) or how to use the score sheet.

Subsequently, they are tested by serving them a subset of the product set, usually in triplicate. These data are evaluated by analysis of variance (ANOVA) and other methods, to determine how consistent and discriminating individual panelists are, as well as the panel as a whole. PanelCheck (http://www.panelcheck.com) and SensoMineR (http://sensominer.free.fr) make it relatively simple and easy to analyze these data. If there are issues, then training continues; however, if all panelists perform to an acceptable standard, then the actual data collection starts.

2.2.3.2 Ballot Training

In a certain sense, consensus training is similar to the panel learning a new language as a child, while ballot training is similar to being taught a language as an adult. In this situation, the panelists do not generate the attributes used to describe differences among the samples but are taught to use an attribute list with reference standards. This attribute list may have been generated as a lexicon for the product category, usually with suggested reference standards included (e.g., Civille and Lyon 1996; Dooley et al. 2009; Lawless et al. 2012; Warmund and Elmore 2008).

Alternatively, the specific attributes (and their reference standards) may have been generated in a consensus training in the first year of a longitudinal study. In this case, the panelists in the second and subsequent years are taught the initial attribute list. It usually takes longer to train a panel using the ballot training method, but there are situations where it may be the only option.

The process for ballot training is similar to consensus training in the sense that in the first session, the panelists are given the two or three most different samples in the product set. They are then asked to use the ballot containing predetermined and defined attributes to describe how the samples differ. During subsequent sessions, this process is repeated until the panel is confident that they understand the attributes, that they can identify the reference standards blind, and that they are consistent in using the attributes. The panel will then be tested in a similar fashion to the consensus-trained panels prior to the actual sample evaluation.

2.2.3.3 Reference Standards

Reference standards have two roles in a DA. First, they anchor the concept assigned to the attributes for the panelists. It is not unusual for two panelists to use different words to describe the same underlying attribute nor is it unusual for two panelists to use the same word to describe different underlying concepts. For example, we had a red wine panel in which most of the panelists said a wine smelled like *Blackberry Jam*, while one panelist insisted it smelled like *Violets*. When the panel leader produced both a *Blackberry Jam* and a *Violet* reference standard, the lone holdout realized that his concept of *Violet* was actually *Blackberry Jam*. We have also had the situation where a number of panelists would agree that a specific sample smelled *Woody*. Yet when the panel leader produced a *Woody* reference standard created by using oak chips in wine, there was intense disagreement. It transpired that for some panelists, *Woody* was actually the aroma associated with the debris found on a forest floor. In this case, the situation was resolved by using one term called *Oak* and another term called *Mushroom*.

Second, reference standards act as translation devices for anyone reading the reports or articles about the study. Lund et al. (2009) used the attribute *Bourbon* to describe differences among Sauvignon blanc wines. On first glance, this term seems nonsensical, until when one realizes that the reference standard used was 1-hexanol, which smells *grassy, chemical*. This is logical since numerous Sauvignon blanc wines are grassy in odor. For this reason, the reference standard recipes should be detailed enough for someone else to recreate them. Tables 2.1 and 2.2 show two reference standard lists—Table 2.1 is inadequate as a translation device and Table 2.2 would be acceptable. Earlier in my career (H. Heymann),

TABLE 2.1
Examples of Reference Standards: Reference Standards for Chocolate Ice Cream Made with Varying Levels of Fat

Attribute	Reference Standard
Color	Light brown to dark brown
Foaminess	Look for bubbly foam
Separation of color	Look for dark and light streaks in melted ice cream
Chocolate	Hershey's™ milk chocolate bar[a] and cocoa used in mix
Cocoa	Cocoa powder and unsweetened chocolate references
Cooked milk aroma	Evaporated milk (Schnucks evaporated milk[b])
Creamy	Combination of thickness and lubricative feeling as ice cream melts—refer to skim milk and cream

Source: Adapted from Prindiville, E.A. et al., *J. Dairy Sci*., 83, 2216, 2000.

[a] Hershey Foods Corporation, Hershey, PA.

[b] Schnucks Foods, St. Louis, MO.

TABLE 2.2
Examples of Reference Standards: Reference Standards for Sauvignon Blanc Wines

Attribute	Reference Standard
Sweet sweaty passion fruit	2000 ng/L 3-mercaptohexyl acetate (Oxford Chemicals)[a]
Bell pepper (capsicum)	1000 ng/L 2-methoxy-3-isobutylpyrazine (Acros Organics)[a]
Cat pee/boxwood	1000 ng/L 4-mercaptomethyl pentane (Oxford Chemicals)[a]
Passion fruit skin/stalk	2000 ng/L 3-mercaptohexan-1-ol (Interchim)[a]
Bourbon	2400 µg/L hexanol/L (Sigma)[a]
Apple candy	250 mg hexyl acetate/L (Sigma)[a]
Tropical	40 mL Golden Circle Mango juice + 40 mL Golden Circle Golden Pash drink + 200 mL Just Juice Mandarin Passion Fruit juice[b]
Mint	25 mg/L cineole (Sigma)[a]
Fresh asparagus	50 mL steamed asparagus water[b]
Stone fruit	Canned Watties apricot and peach juice soaked in diluted base wine for 30 min (equal parts)[b]

Source: Adapted from Lund, C.M. et al., *Am. J. Viticult. Enol*., 60, 1, 2009.

[a] Added to diluted base wine (50% Corban Sauvignon blanc and 50% water).

[b] Added equal parts to base wine (Corban Sauvignon blanc).

I published a few papers (e.g., Im et al. 1994; Lin et al. 1998) without reference standards. However, once I realized how important they are to translations, I have tried to make the tables with reference standards as clear as possible.

Reference standards could be made using chemicals, for example, Lund et al. (2009) used 1000 ng/L mercaptomethyl pentane in a diluted base wine (50% Sauvignon blanc wine and 50% water) as a reference standard for cat pee/boxwood. However, in many places, due to environmental health and safety rules, sensory laboratories are not allowed to use chemicals since food and beverages are evaluated in that space. In these cases, it is easier and often the only legal option to use actual products to simulate the required concepts, for example, Robinson et al. (2011) cut 2 cm lengths of leather shoelaces (Kiwi Outdoor shoelaces) into small squares and then soaked them in 50 mL base red wine (Franzia Vintners Select Cabernet Sauvignon) as the reference standard for the *leather* aroma in red wines (Robinson et al. 2011).

As a last resort, the reference standard could be anchored by a verbal definition. This truly should be a rare occurrence since this type of standard is neither good for concept anchoring nor good as a translation device. One of the few times, recently, that we have used verbal descriptors was in a study of chocolate milks, where a few of the milks had a fecal off-odor. The panel leader created a reference standard by scraping fecal matter from the floor of a cow barn. The panelists, after smelling it once, decided that they did not need to smell this reference again, and for the remainder of the study, the attribute was verbally anchored.

Reference standard creation is part science and part art. The most complex part is to determine exactly what the panel means when they say a specific word. For example, in a recent Chardonnay study, the panel wanted an *Apple* reference standard. The panel leader created a number of potential apple standards (Table 2.3) and then asked each panelist to score each standard on a 1–9-point numerical scale in terms of its match to their mental concept of *Apple* as they perceived it in the samples under discussion. A score of 1 was assigned when the reference standard was an exact match and a 9 was assigned when the standard had no relationship to the concept. From this, a median score can be calculated and it is fairly easy to determine which standard should be used. In the case of the Chardonnay panel, Apple 5 was the closest match to the concept of *Apple*. We use this technique for all our reference standards. A similar technique, using an appropriateness scale, has successfully been used by Murray and Delahunty (2000).

Panelists should be able to identify the reference standards blind. This is accomplished by giving them a list of attributes and a set of

TABLE 2.3
Potential Apple Reference Standards for Chardonnay Wines

	Reference Standard	Median[a]
Apple 1	20 g Red Delicious fresh apple, chopped + 25 mL base wine	5
Apple 2	20 g Red Delicious fresh apple, chopped + 25 mL base wine; decanted after 1 h; serve liquid as standard	4
Apple 3	8 g Granny Smith fresh apple, chopped + 25 mL base wine	6.5
Apple 4	8 g Granny Smith fresh apple (in one piece) + 25 mL base wine	4.5
Apple 5	12.5 g Granny Smith fresh apple, chopped + 25 mL wine	1.5[b]
Apple 6	8 g canned Pie Fruit Apples, sliced (Ardmona, Victoria, Australia) + 25 mL base wine	5
Apple 7	10 g canned Granny Smith Apple Slices (WW Select, Woolworths, Australia) + 25 mL base wine	4.5
Apple 8	Orchard Apple Stage 1 Baby Food (only organic, Auckland, New Zealand) + 25 mL wine	4.5
Apple 9	25 mL 100% Granny Smith cold-pressed juice (Preshafruit, Victoria, Australia) + 25 mL base wine	2.5
Base wine	Sunnyvale Dry White Wine, Miranda Wines, Merbein, Victoria, Australia	

[a] Mean score based on 1 = reference standard very similar to the concept of *Apple* in these wines and 9 = reference standard very dissimilar to the concept of *Apple* in these wines.
[b] Standard used in the actual study.

reference standards labeled by three-digit codes. The panelists are asked to match the code of the reference standard to the correct attribute. Once the panelists can identify all standards consistently and accurately, the data collection phase can start. This process is repeated as a reference standard identification test in the booths prior to each evaluation session to ensure that panelists interact with the reference standards. In the case of a computerized data acquisitions system, it is fairly easy to do this and to provide panelists with instant feedback on their accuracy. After the first few sessions, it is extremely rare for panelists to identify the standards incorrectly. If a panelist's performance suddenly drops, it indicates to the panel leader that there may be an issue that needs to be explored.

2.2.4 Evaluation of Samples

Once the panel has been trained and tested, then the actual evaluation of the samples can commence. It is usual that this process occurs in individual

temperature- and light-controlled booths (Lawless and Heymann 2010), but it is also possible to do the evaluation in a large conference room, as long as the panelists are not within each other's line of sight and there is no discussion or distractions (Snitkjaer et al. 2011).

These days, data acquisition is usually performed through a computerized system (e.g., Compusense, Guelph, Canada; EyeQuestion, Elst, the Netherlands; and FIZZ, Couternon, France), but the use of paper ballots is not unusual. There is an indication that switching from paper ballots to computerized ballots is not detrimental to the data collection (Swaney-Stueve and Heymann 2002), and in some cases, this is helpful, for example, when a computer glitch prevents the use of the computerized acquisition system but the samples have already been prepared.

Panelists must be made to feel welcome and appreciated during the data acquisition phase to ensure continued motivation and interest. It is not unusual to serve them some snacks as a token of appreciation after they complete their sensory sessions. In certain situations, it may also be appropriate to pay panelists.

2.2.5 Data Analysis

The next chapter in this book is on multivariate data analysis, and thus, we will not provide an in-depth discussion in this chapter. However, it is beneficial to describe the standard sequence in which we start the data analysis process in our laboratory. Assuming that we had a fairly uncomplicated experimental design involving samples, panelists, and replications and that we have no missing values,* we start with a three-way multivariate analysis of variance (MANOVA) with a related series of univariate analyses (ANOVA) of all attributes. In this case, the main effects would be samples, panelists, and replication with the addition of all two-way interactions (panelists by sample, panelist by replication, and sample by replication). The MANOVA tests for the overall significance of all the attributes in the data and the ANOVAs for the individual attributes.

* If there are missing values, for example, where a panelist missed a session, that could lead to complications with multivariate data analysis techniques. For this reason, if the number of missing values is less than 10% (and it is usually 2% or less), we usually impute the missing variables by calculating the average of the two (out of three) replication that the panelist actually evaluated. This decreases the overall variability of the data, and thus, the analyst should remove an equivalent number of degrees of freedom from the error or residual term in the MANOVA and the individual ANOVAs (Beale and Little 1975; Little and Rubin 1987).

If the MANOVA is significant with a probability of 5% or less, then we continue and evaluate the significance levels of the individual attributes. If sample is significant, as well as the sample by panelist or the sample by replication interaction terms is, then we need to evaluate the impact of this interaction on the sample effect. The standard in our laboratory is to use the pseudomixed model (Naes and Langsrud 1998) where the F-value for sample is calculated by dividing the mean square (sample) value by either the mean square (sample by panelist) value or the mean square (sample by replication) value. If the calculated sample F-value remains significant, then we assume that the interaction effect is not important and we treat that attribute as significant for the sample effect. If the F-value is not significant, then the interaction has an impact on the sample effect and we treat that attribute as not significant for the sample effect. There are other ways in which these data could be analyzed and we suggest the following references: Lawless (1998), Schlich (1998), and Steinsholt (1998).

For any significant attributes, we would then calculate a mean separation value for the means of the samples for each attribute. We traditionally use Fisher's protected least significant difference (LSD), but these values are somewhat liberal, and if a more conservative value is needed, we would use Tukey's honestly significant difference (HSD). See Gacula et al. (2008) for further discussion on mean separation techniques.

The next step of the data analysis involves a graphical representation of the data. Our preference is the creation of a canonical variate analysis (CVA). To do this, rerun the MANOVA (main effect: wine) since Monrozier and Danzart (2001) have shown that the one-way analysis is more stable in calculating a CVA. We use CVA as a multivariate *mean separation* technique for the MANOVA (Chatfield and Collins 1980). The CVA will separate the mean positions of the samples in a 2D or 3D space. It is possible to calculate the number of significantly discriminating dimensions using Bartlett's test (Bartlett 1947; Chatfield and Collins 1980) as well as the 95% confidence intervals around mean position of each sample (Chatfield and Collins 1980; Owen and Chmielewski 1985). These pieces of information make the CVA more useful than the principal component analysis (PCA). For further discussion on the advantages of CVA over PCA, see Heymann and Noble (1989) and Monrozier and Danzart (2001).

2.3 CASE STUDIES

In this section, we discuss a fairly simple case study involving six commercial Cabernet Sauvignon wines and blends of Cabernet Sauvignon wines. We then discuss a more complex study involving 17 commercial wines from 6 countries. The data sets are available for download from the CRC Web site: http://www.crcpress.com/product/isbn/9781466566293.

TABLE 2.4
Commercial Cabernet Sauvignon Wines and Blends for Case Study 1

Code	Vintage	Blend[a]	Wine Region and/or State	Retail Price (US$)	Alcohol Content (%v/v)
W1	2009	100% CS	Paso Robles, California	6	13.2
W2	2008	88% CS, 10% CF, 2% Merlot	Napa Valley, California	42	14.0
W3	2006	100% CS	Napa Valley, California	68	15.2
W4	2006	100% CS	Napa Valley, California	60	15.2
W5	2007	60% CS, 15% Syrah, 11% Merlot, 10% Petit Verdot, 4% CF	Columbia Valley, Washington State	26	15.5
W6	2008	100% CS	Washington State	50	15.9

[a] CS, Cabernet Sauvignon; CF, Cabernet Franc.

2.3.1 Case Study 1

Six commercial wines (Table 2.4) made with at least 60% Cabernet Sauvignon were evaluated in quadruplicate by 11 trained panelists.* The panel had been trained over five sessions using the consensus method sequence described in Section 2.2.3.1. The panel used 12 attributes (Table 2.5) to describe differences among the wines.

The data were analyzed using R and all R-code is shown in Appendix 2.A. A MANOVA (main effects: panelists, wines, replications, and all two-way interactions; Table 2.6) was followed by a series of ANOVAs (main effects and interactions as in the MANOVA; Figure 2.1) where the pseudo-mixed model was used whenever the wine interactions (wine by panelists and/or wine by replication) were significant. This was the case for HerbalA (herbal aroma), AlcoholA (alcohol aroma), and BurningA (burning aroma), where the former two attributes remained significant after the application of the pseudomixed model and the last one became nonsignificant.

* These data are related to, but not part of, the study described in King et al. (in press). The quadruplicate analysis of each sample was an artifact of the specific study and is not the usual way we do replication. Triplicates are more standard.

TABLE 2.5
Attributes and Reference Standards Used for Case Study 1

Attribute	Description	Reference Standard
Aroma (A)		
Fresh fruit	Red apple, banana, orange, peach, pear, pomegranate, grape, mango, citrus	2 pieces red and yellow papaya from canned tropical fruit (Dole), 1/2 cm^2 piece fresh banana, 1/2 cm^2 piece fresh apple, and 1/2 cm piece fresh lemon rind
Berry	Blackberry, blueberry, raspberry, strawberry, tart berry, forest fruit	1 fresh strawberry halved, 1 fresh raspberry halved, and 1 fresh blackberry halved
Herbal	Grassy, leafy	1 tsp fresh, cut grass and 1 tsp of green leaves
Barnyard	Brett, band-aid	1 grain 4-ethylphenol
Alcohol		1 tsp Vodka (Sobieski)
Burning	Physical prickling sensation in nose	—
Taste and mouthfeel (T)		
Sourness	Acidity, tart	2 g/L tartaric acid (Fisher Scientific) dissolved in water
Sweetness		15 g/L (D)-fructose (Sigma) dissolved in water
Bitterness		800 mg/L anhydrous caffeine (Sigma) dissolved in water
Alcohol	Warm to hot	150 mL/L Vodka (Sobieski) in water
Viscosity	Thickness of mouthfeel, body of wine, oiliness Low anchor (thin) High anchor (thick)	7 g/L Pectin ex-citrus (Sigma) dissolved in water
Astringency	Dry, tannic, puckering	800 mg/L alum (McCormick) dissolved in water

Source: Adapted from King, E.S. et al., *Am. J. Enol. Viticult*., 2012.

Fisher's LSD was used as a univariate mean separation technique for all attributes that differed significantly across wines (Table 2.7). A one-way MANOVA with wine as the main effect was followed by a CVA used as a multivariate *mean separations technique* (Figure 2.2). Additionally, a PCA was performed on the covariance matrix of the mean wine values, shown in Figure 2.3.

TABLE 2.6
MANOVA Table Obtained from R (See R-Code in Appendix 2.A) Using a Three-Way Fixed-Effect Model with All Two-Way Interactions for Case Study 1

SoV[a]	Degrees of Freedom (df)	Wilk's Lambda	Approximate F-Value	Numerator df	Denominator df	Probability (P) > F[b]
Wine	5	0.32832	2.9306	60	654.66	2.21e−11*
rep[c]	3	0.61819	2.0203	36	411.42	6.379e−4*
pan[c]	10	0.00104	12.9762	120	1093.52	<2.2e−16*
w/r[c]	15	0.27444	1.0899	180	1335.53	0.2108887
w/p[c]	50	0.01492	1.2056	600	1669.02	2.345e−3*
r/p[c]	30	0.01358	2.0751	360	1591.18	<2.2e−16*
res[c]	150					

[a] SoV, sources of variation.

[b] * indicates 0.05.

[c] rep, replication; pan, panelist; w/r, wine/replication; w/p, wine/panelist; r/p, replication/panelist; res, residuals.

		Fresh Fruit Aroma (FrshFrtA)				Berry Aroma (BerryA)			
SoV[a]	**df[a]**	**SS[a]**	**MS[a]**	**F-Value**	**Pr(>F)**	**SS[a]**	**MS[a]**	**F-Value**	**Pr(>F)[b]**
wine	5	80.06	16.01	3.01	0.01*	86.60	17.32	3.61	0.00*
rep[c]	3	6.44	2.15	0.40	0.75	29.23	9.74	2.03	0.11
pan[c]	10	310.7	31.07	5.84	0.00*	169.6	16.96	3.53	0.00*
w:r[c]	15	33.32	2.22	0.42	0.97	89.56	5.97	1.24	0.24
w:p[c]	50	227.0	4.54	0.85	0.74	290.2	5.80	1.21	0.19
r:p[c]	30	252.5	8.42	1.58	0.04*	350.5	11.68	2.43	0.00*
res[c]	150	798.0	5.32			719.5	4.80		
		Herbal Aroma (HerbalA)				**Barnyard Aroma (BrnYrdA)**			
wine	5	38.39	7.67	3.86	0.00*	36.82	7.36	2.39	0.04*
Wine[d]	5			2.58	0.037*				
rep[c]	3	1.15	0.38	0.19	0.90	2.58	0.86	0.28	0.84
pan[c]	10	324.4	32.44	16.3	0.00*	455.6	45.56	14.8	0.00*
w:r[c]	15	50.64	3.38	1.70	0.06	34.15	2.277	0.74	0.74
w:p[c]	50	148.6	2.97	1.49	0.03*	212.6	4.25	1.38	0.07
r:p[c]	30	152.8	5.09	2.56	0.00*	94.91	3.16	1.02	0.43
res[c]	150	298.5	1.99			461.5	3.08		
		Alcohol Aroma (AlcoholA)				**Burning Aroma (BurningA)**			
wine	5	54.36	10.87	6.71	0.00*	18.66	3.73	2.74	0.02*
Wine[e]	5			4.03	0.00*			1.40	0.24
rep[c]	3	14.16	4.72	2.91	0.04*	9.11	3.04	2.23	0.09
pan[c]	10	624.2	62.42	38.5	0.00*	430.5	43.05	31.6	0.00*
w:r[c]	15	37.61	2.51	1.55	0.09	38.14	2.54	1.87	0.03*
w:p[c]	50	134.8	2.70	1.66	0.01*	133.5	2.67	1.96	0.00*
r:p[c]	30	81.47	2.72	1.67	0.02*	59.99	2.00	1.47	0.07
res[c]	150	243.1	1.62			204.3	1.36		
		Sour Taste (SourT)				**Sweet Taste (SweetT)**			
wine	5	25.13	5.03	1.46	0.20	17.61	3.52	1.08	0.37
rep[c]	3	9.56	3.19	0.92	0.43	1.85	0.62	0.18	0.90
pan[c]	10	422.4	42.24	12.25	0.00*	581.7	58.17	17.81	0.00*
w:r[c]	15	38.48	2.56	0.74	0.73	49.35	3.29	1.01	0.45
w:p[c]	50	156.1	3.12	0.90	0.65	116.7	2.33	0.71	0.91
r:p[c]	30	214.6	7.15	2.07	0.00*	236.9	7.90	2.42	0.00*
res[c]	150	517.2	3.45			490.0			

FIGURE 2.1 ANOVA tables for all attributes evaluated in for Case Study 1. See R-codes in Appendix 2.A.

(*continued*)

		Bitter Taste (BitterT)				Alcohol Mouthfeel (AlcoholT)			
wine	5	26.17	5.23	1.92	0.09	33.31	6.66	2.99	0.01*
rep[c]	3	36.91	12.30	4.51	0.00*	30.12	10.04	4.51	0.00*
pan[c]	10	464.3	46.43	17.02	0.00*	764.4	76.4	34.32	0.00*
w:r[c]	15	53.22	3.55	1.30	0.21	32.53	2.17	0.97	0.48
w:p[c]	50	95.96	1.92	0.70	0.92	95.05	1.90	0.85	0.74
r:p[c]	30	231.7	7.72	2.83	0.00*	99.66	3.32	1.49	0.06
res[c]	150	409.0	2.73			334.1	2.23		
		Viscous Mouthfeel (ViscosityT)				Astringent Mouthfeel (AstringencyT)			
wine	5	28.36	5.67	2.04	0.07	149.5	29.90	10.01	0.00*
rep[c]	3	9.68	3.23	1.16	0.32	18.60	6.20	2.08	0.11
pan[c]	10	157.9	15.79	5.69	0.00*	410.7	41.07	13.75	0.00*
w:r[c]	15	64.97	4.33	1.56	0.09	34.86	2.32	0.78	0.79
w:p[c]	50	188.0	3.76	1.36	0.08	154.8	3.10	1.04	0.42
r:p[c]	30	257.5	8.58	3.09	0.00*	119.1	3.97	1.33	0.13
res[c]	150	416.1	2.77			447.9	2.99		

[a] SoV, sources of variation; df, degrees of freedom; SS, sums of squares; MS, mean sums of squares; Pr(>F), probability larger than F-value.

[b] * indicates 0.05.

[c] rep, replication; pan, panelist; w:r, wine:replication; w:p, wine:panelist; r:p, replication:panelist; res, residuals.

[d] The shaded wine row lists the pseudomixed model for Herbal Aroma (7.67/2.97 = 2.58) with numerator df = 5 and denominator df = 50.

[e] The shaded wine row lists the pseudomixed model for Alcohol Aroma (10.87/2.70 = 4.03) with numerator df = 5 and denominator df = 50 and the pseudomixed model for Burning Aroma (3.73/2.67 = 1.40) using the w:p interaction (since it was larger than the w:r interaction) with numerator df = 5 and denominator df = 50.

FIGURE 2.1 (continued) ANOVA tables for all attributes evaluated in for Case Study 1. See R-codes in Appendix 2.A.

The CVA (Figure 2.2) shows that the first two dimensions (both of which were significant) explain a total of 85.5% of the variance ratio in the data space. The 95% confidence ellipse of W1 does not overlap any of the other wines' confidence ellipses. This wine is significantly different from all the other wines, and when we look at the means table (Table 2.7), we find that W1 was higher in Fresh Fruit and berry aromas than all the wines except W4 (for Fresh Fruit). Additionally, W1 had the lowest perceived alcohol aroma and flavor (Table 2.7), and as can be seen in Table 2.4, W1 had the lowest alcohol content as well. According to the CVA (Figure 2.2), W2 is significantly different from W4, W5, and W6 but not from W3. Table 2.7 indicates that W5 and W6 were significantly more astringent in

TABLE 2.7
Means and Fisher's LSD Values for the Significantly Different Attributes for Wines in Case Study 1

Attribute[a]	W1	W2	W3	W4	W5	W6	LSD[b]
FrshFrtA	3.5 a	2.0 bc	2.3 bc	2.9 ab	1.9 c	2.5 bc	1.0
BerryA	4.3 a	2.5 b	3.3 b	3.0 b	2.7 b	3.0 b	0.9
HerbalA	2.2 a	1.4 bc	0.9 c	1.7 ab	1.5 bc	1.2 bc	0.6
BarnYrdA	2.2 ab	2.4 a	1.8 abc	1.6 bc	1.3 c	1.6 bc	0.7
AlcoholA	1.2 c	2.3 ab	1.8 b	2.3 ab	1.8 b	2.6 a	0.5
AlcoholT	2.2 b	2.8 a	3.1 a	2.8 ab	3.1 a	3.2 a	0.6
AstringencyT	2.1 c	2.6 c	3.4 b	3.4 b	4.2 a	4.1 a	0.7

Note: See R-code in Appendix 2.A.

[a] FrshFrtA, Fresh Fruit aroma; BerryA, berry aroma; HerbalA, herbal aroma; BarnYrdA, barnyard aroma; AlcoholA, alcohol aroma; AlcoholT, alcohol mouthfeel; AstringencyT, astringent mouthfeel.

[b] Means with the same letter within a row are not significantly different.

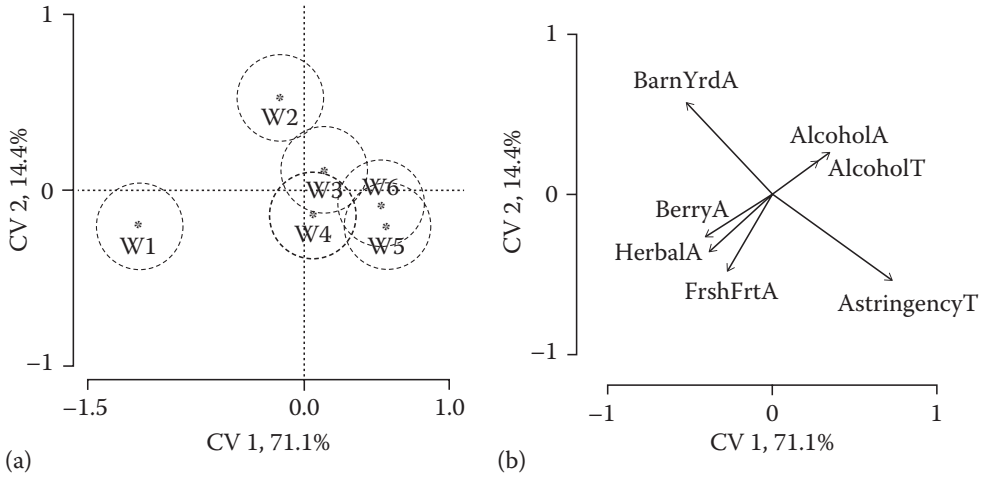

FIGURE 2.2 CVA score (a) and loading (b) plots of the significantly different sensory attributes for six commercial Cabernet Sauvignon wines and blends evaluated by 11 panelists in quadruplicate in Case Study 1. The circles represent 95% confidence intervals and circles that overlap indicate wines that are not significantly different.

mouthfeel than any of the other wines and this can also be inferred from their positions on the CVA. In this case study, the CVA (Figure 2.2) and the PCA (Figure 2.3) are very similar but inverted. The PCA explained 83.8% of the variance in the data space, which is slightly less than the first two dimensions of the CVA.

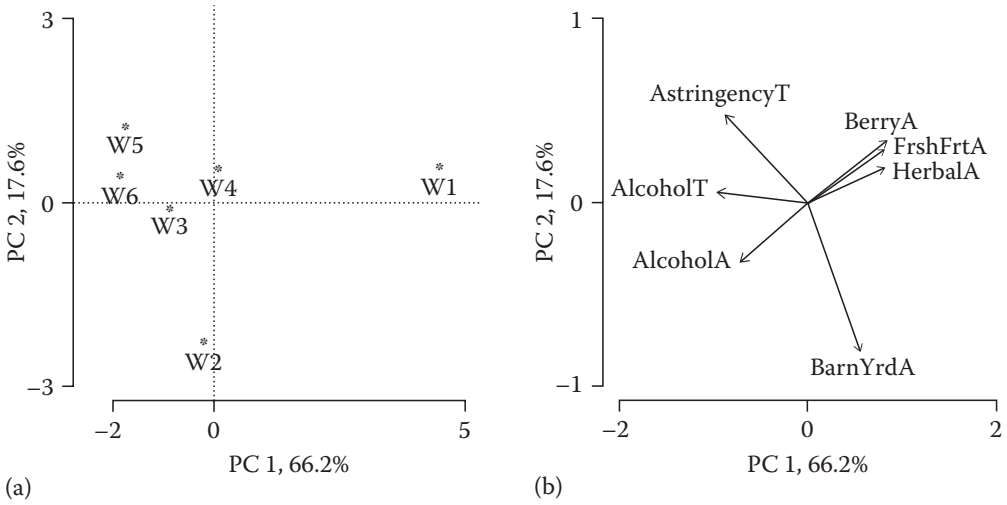

FIGURE 2.3 PCA score (a) and loading (b) plots of the significantly different sensory attributes for six commercial Cabernet Sauvignon wines and blends for six commercial Cabernet Sauvignon wines evaluated by 11 panelists in quadruplicate in Case Study 1.

2.3.2 Case Study 2

In this study, 17 red wines from 6 countries were evaluated by 22 panelists* using 18 attributes (Machado 2009). The wines are listed in Table 2.8 and the attributes and reference standards are shown in Table 2.9.† The data were analyzed using R and all R-code is shown in Appendix 2.B. The data analysis process was very similar to Case Study 1. The R-code is shown in Appendix 2.B. The MANOVA was significant for wines (data not shown), and we then did univariate ANOVAs (data not shown) on all the attributes; after evaluating the effects of significant wine panelist interactions on the wine main effect, the following 12 attributes were significantly different at $P < 0.05$ for the following wines: AlcoholA (alcohol aroma), CitrusA (citrus aroma), VeggieA (veggie aroma), CaramelA (caramel aroma), WoodyA (woody aroma), LeatherA (leather aroma), MedicinalA (medicinal aroma), SweetT (sweet taste), BitterT (bitter taste), BodyVisT (viscous mouthfeel), and AStrinT (astringent mouthfeel).

These attributes were used in the CVA (Figure 2.4). The first three dimensions of the CVA were significant and the first two explained 56.3%

* This study had an unusually large number of panelists on the panel. The reason was that we were collecting data for a project trying to determine the optimum number of panelists in a descriptive analysis study (Heymann et al. 2012).

† To save space, only the attributes that were significantly different across wines are shown.

TABLE 2.8
Commercial Red Wines from Six Countries Used for Case Study 2

Wine	Varietal Blend	Appellation	Alcohol (%v/v)	Price (US$)
AR1	Malbec	Mendoza, Argentina	13.5	49.99
AR2	65% Malbec, 35% Cabernet Sauvignon	Mendoza, Argentina	14.0	42.99
AU1	Shiraz	Barossa Valley, South Australia	14.5	88.99
AU2	Shiraz	Barossa Valley, South Australia	14.8	65.00
CH1	73% Cabernet Sauvignon, 22%–23% Carmenere, 4%–5% Cabernet Franc	Chile	14.5	64.99
FR1	Bordeaux Blend	Pomerol, Bordeaux	14.0	73.00
FR2	52% Cabernet Sauvignon, 45% Merlot, 3% Petit Verdot	Margaux, Bordeaux	13.5	84.99
FR3	55% Merlot, 45% Cabernet Sauvignon	Pessac Leognan, Bordeaux	13.5	191.99
PT1	Portuguese Blend	Alentejo, Portugal	14.1	16.00
PT2	Portuguese Blend	Dão, Portugal	14.5	16.00
PT3	Portuguese Blend	Douro, Portugal	12.5	16.00
PT4	Portuguese Blend	Alentejo, Portugal	14.0	25.00
PT5	Portuguese Blend	Dão, Portugal	15.4	25.00
PT6	Portuguese Blend (Tinta Roriz, Touriga Nacional, Touriga Franca)	Douro, Portugal	12.2	25.00
PT7	Portuguese Blend	Douro, Portugal	14.0	40.00
US1	Cabernet Sauvignon	Napa Valley, United States	14.9	76.99
US2	Syrah	Napa Valley, United States	14.8	54.99

TABLE 2.9
Attributes and Reference Standards for Significant Attributes Used for Case Study 2

Aroma Attributes	Code	Composition
Alcohol	AlcoholA	15 mL of ethanol solution (25 mL 100% ethanol/475 mL water)
Caramel	CaramelA	1/2 tbs soy sauce + 1/2 tbs molasses + 1/2 tbs butter
Citrus	CitrusA	1 tsp of Bigelow Tea Earl Grey Tea (R.C. Bigelow Fairfield, CT 06825) + 1/2 tsp orange zest + 1/2 tsp lemon zest
Leather	LeatherA	4 × 1/6 in. pieces of Leather Kiwi Outdoor yellow/brown shoelaces
Medicinal	MedicinA	1 drop of 4-ethylphenol solution (1 g/L)
Veggie	VeggieA	10 mL canned Green Bean Brine (Del Monte Fresh cut blue lake French style Green Beans 411 g) + 10 mL canned Asparagus Brine (Raley's fine foods Whole Asparagus Spears 425 g)
Woody	WoodyA	3 chips (evOAK, high vanilla) + 1 drop of Wright's All Natural Hickory Seasoning Liquid Smoke
Taste and mouthfeel attributes		
Astringent	AstrinT	312 mg alum/500 mL water
Bitter taste	BitterT	800 mg caffeine/500 mL water
Sweet taste	SweetT	20 g sucrose/500 mL water
Viscosity/body	BodyVisT	2.5 g pectin/500 mL water
Sour taste	SourT	200 mg citric acid/500 mL water

Source: Adapted from Machado, B., Revealing the secret preferences for top-rated dry red wines through sensometrics, MS thesis, University of California at Davis, Davis, CA (Advisor: H. Heymann), 2009.

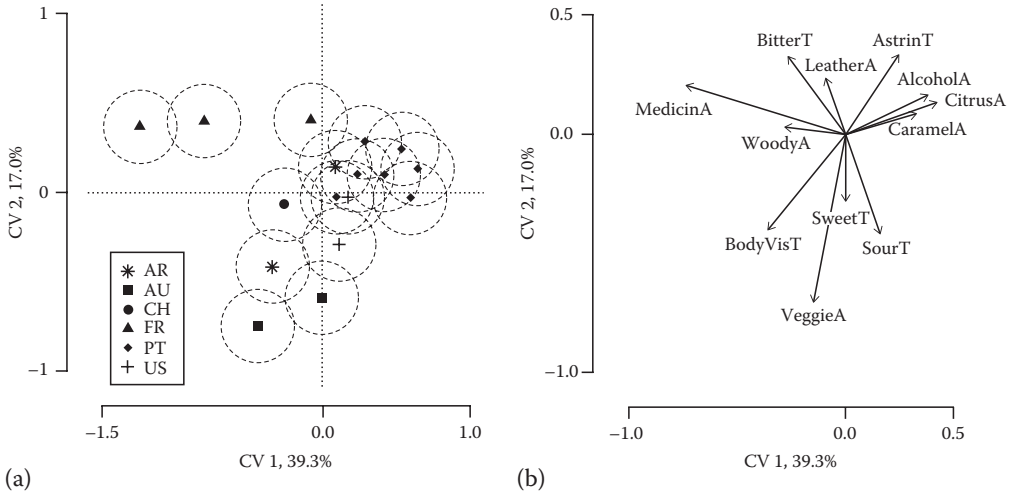

FIGURE 2.4 CVA score (a) and loading (b) plots of the significantly different sensory attributes for 17 commercial red wines from six countries evaluated by 22 panelists in triplicate in Case Study 2. The circles represent 95% confidence intervals and circles that overlap indicate wines that are not significantly different. (See R-codes in Appendix 2.B.)

of the variance ratio. The percentage of the variance ratio explained is much lower than in Case Study 1, but this is not unexpected since there were more wines. The Australian wines were not different from each other, and they differed from all the other wines except an American and an Argentine wine. From the CVA, it would seem that the Australian wines were more vegetative and viscous. The means table (Table 2.10) shows that one of the Australian wines is the most vegetative and that they are both very viscous. The French wines differed from all others except an American and a Portuguese wine. From the CVA, it would seem that the French wines were medicinal and leathery. Table 2.10 indicates that two of the French wines are very high in medicinal and leather intensities. The Portuguese wines clustered together and did not differ from one another.

2.4 CONCLUSIONS

In this chapter, we described classical DA, highlighting the two ways to generate attributes (consensus and ballot training). We also described the intricacies of reference standards. In the final section, we very briefly showed two case studies, with their R-code. The intent was to make sure that the reader has a thorough understanding of the classical DA methodology before plunging into the remainder of the book where more novel techniques in profiling and sensory science will be highlighted.

TABLE 2.10
Means and Fisher's LSD Values for Significantly Different Attributes for Case Study 2

Wine	CitrusA	AlcoholA	VeggieA	CaramelA	WoodyA	LeatherA
AR1	0.7 ef	2.5 d	1.8 b	1.2 ef	3.2 cdef	1.6 cdef
AR2	0.9 ef	2.5 d	0.9 ef	1.6 abcde	2.7 fg	1.3 defgh
AU1	0.7 ef	2.9 bcd	2.6 a	1.2 ef	3.3 bcd	1.9 bc
AU2	0.8 def	3.2 ab	1.5 bcd	1.4 bcde	3.4 bcd	1.1 gh
CH1	0.7 f	3.1 abc	1.6 bc	1.2 ef	3.6 abc	1.6 cde
FR1	0.6 f	2.9 bcd	1.0 ef	1.5 abcde	3.4 bc	1.7 cd
FR2	0.6 f	2.7 cd	1.4 bcde	1.3 cdef	4.0 a	2.1 ab
FR3	0.6 f	2.6 d	1.5 bcd	0.9 f	4.0 a	2.4 a
PT1	0.8 def	2.8 bcd	1.1 def	1.9 ab	2.8 fg	1.5 defg
PT2	1.7 a	2.9 abcd	0.9 ef	1.8 abc	2.8 efg	1.2 fgh
PT3	1.4 abc	2.5 d	1.2 cde	1.3 def	2.4 g	0.9 h
PT4	0.7 f	2.8 bcd	1.4 bcd	1.4 cde	3.4 bcd	1.2 fgh
PT5	1.1 bcd	3.1 abc	1.0 ef	1.7 abcd	3.3 bcde	1.2 efgh
PT6	1.5 ab	2.5 d	0.7 f	1.2 ef	2.9 def	1.0 h
PT7	0.9 def	2.6 d	0.7 f	1.3 def	3.3 bcde	1.3 efgh
US1	0.6 f	3.1 abc	1.6 bc	1.9 a	3.7 ab	1.3 defgh
US2	1.1 cde	3.4 a	1.2 cde	1.3 cdef	2.9 def	1.4 defgh
LSD	0.4	0.5	0.5	0.5	0.5	0.4

Wine	MedicinalA	SweetT	SourT	BitterT	BodyVisT	AstrinT
AR1	1.3 c	1.3 cde	2.7 abcdef	1.6 ef	4.0 bc	3.1 c
AR2	1.1 cd	1.3 cd	2.6 bcdef	2.0 cdef	3.7 c	3.8 cd
AU1	1.2 cd	1.7 b	2.4 ef	1.8 def	4.5 a	3.3 de
AU2	0.9 cde	2.1 a	2.9 ab	1.6 ef	4.5 a	3.3 de
CH1	1.2 c	1.3 cdef	2.8 abcde	2.2 abcd	4.0 bc	4.0 abc
FR1	1.0 cde	0.9 h	2.5 cdef	2.5 ab	3.8 c	4.3 ab
FR2	2.0 b	1.0 fgh	2.5 def	2.7 a	4.0 bc	4.3 abc
FR3	2.9 a	1.0 gh	2.6 bcdef	2.2 abcd	3.9 bc	4.4 a
PT1	0.8 de	1.0 defgh	2.8 abcdef	2.0 bcde	4.1 bc	4.2 abc
PT2	0.6 e	1.2 cdefg	2.6 bcdef	1.7 ef	3.9 bc	4.1 abc
PT3	0.6 e	0.9 h	2.9 abc	1.5 f	2.9 d	3.2 e
PT4	0.7 de	1.0 efgh	3.1 a	2.3 abc	3.8 c	4.4 a
PT5	0.6 e	1.3 cdefg	2.6 bcdef	1.7 ef	4.0 bc	4.2 abc
PT6	0.7 de	0.9 h	2.9 abcd	1.5 f	3.0 d	3.5 de
PT7	0.6 e	1.1 cdefgh	2.4 f	1.7 ef	3.6 c	4.2 abc
US1	0.6 e	1.3 c	2.9 abcde	2.0 cdef	4.2 ab	3.8 bcd
US2	1.1 cd	1.2 cdefg	3.0 a	1.9 cdef	4.3 ab	4.2 abc
LSD	0.4	0.3	0.4	0.5	0.4	0.5

Note: See R-code in Appendix 2.B.

2.A APPENDIX: R-CODE FOR THE CASE STUDY 1

```
# classical Descriptive Analysis evaluation – example 1
# (c) H. Hopfer, October 2012
# all code comes without any warranty

## ---------------read data into R----------------- ##
# in this data set we have 11 judges, 4 replicates,
# 6 wines & 12 sensory attributes

da.d = read.table('data1.csv', sep=',', header=TRUE)
head(da.d)
dim(da.d)

# check your data: 11*4*6 = 264 observations;
# 12 + 3 = 15 columns

# define judge, rep and wine as factor
# columns starting with a letter are automatically set
# as factors by R

for(i in 1:3) {da.d[,i] <- as.factor(da.d[,i])
         print(is.factor(da.d[,i]))}

# combine all attributes and define it as a matrix

da.a = as.matrix(da.d[,-c(1:3)])

# use all columns but the first three (wine, judge, rep)

head(da.a)

## --------------MANOVA--------------- ##
# building a 3-way MANOVA model with all 2-way
# interactions, and run the MANOVA

da.lm = lm(da.a ~ (wine + rep + judge)^2 , data= da.d)
da.maov = manova(da.lm)
summary(da.maov, test='Wilks')

# in MANOVA wine, rep and judge and the interaction
# wine:judge and rep:judge are significant
# at p< 0.05 => continue with individual ANOVA's

## --------------ANOVA---------------- ##
# using the same lm model but now use the ANOVA output

da.aov = aov(da.lm)
aovsum = summary(da.aov)
aovsum

# sign. wine effect at p< 0.05 for: FrshFrtA, BerryA,
# HerbalA, BarnYrdA, AlcoholA, BurningA, AlcoholT,
# AstringencyT
```

```
# sign. wine interaction (p<0.05) with sign. wine
# effect: HerbalA (W:J), AlcoholA (W:J), BurningA
# (W:J,W:R)

## -------------Pseudomixed model--------------- ##
# apply pseudomixed model for attributes with sign.
# wine interactions
# Fnew = (MS_wine)/(MS_wine interaction)
# determine critical F-value for wine effect and W:J
# interaction, df1= df wine and df2 = df W:J

df_W = aovsum[[1]][1,1]
df_WJ = aovsum[[1]][5,1]
newF_crit = qf(0.95, df_W, df_WJ)

# HerbalA => sign. W:J; calculate new F-value and test
# significance for

newF_herbal = aovsum[[3]][1,3]/(aovsum[[3]][5,3])
newF_herbal > newF_crit

# F remains significant at p<0.05 => HerbalA has a
# sign. wine effect

# ditto for AlcA and BurningA

# continue further analyses with only significant
# attributes (significant wine effect)
# create data subset with significant attributes only
# from original data
# sign. attributes are FrshFrtA, BerryA, HerbalA,
# BarnYrdA, AlcA, AlcT, AstrT

da.s = da.d[, c(1:8,13:15)]
head(da.s)
da.s.a = as.matrix(da.s[,-c(1:3)])
head(da.s.a)

## -----------------LSDs----------------- ##
# calculate LSDs or HSDs using the agricolae package
# install the agricolae package first from the CRAN
# and then load it

library(agricolae)

# for LSD use LSD.test, individually for all your
# attributes of interest => see also ?LSD.test

FrshFrtA.lm = lm(FrshFrtA ~ (wine + rep + judge)^2,
  data=da.s)

# build lm models for each attr
```

```
FrshFrtA.LSD = LSD.test(FrshFrtA.lm, trt='wine',
 group = TRUE)

# calculate means and LSD

# ditto for all other significant attributes
# for HSD use HSD.test, analog to LSD.test => see also
# ?HSD.test

## --------------CVA------------------ ##
# run CVA on the significant data set using the
# candisc package
# install candisc package first from the CRAN and load
# it afterwards

library(candisc)

# build MANOVA model with significant attributes only,
# using only the wine effect
# see Monrozier, R.; Danzart, M. (2001) Food Quality
# and Preference 12:393-406

da.s.mlm = lm(da.s.a ~ wine, data = da.s)
da.cva = candisc(da.s.mlm)

# extract CVA output => eigenvalues, variance ratios
# and Bartlett's test for sign. CVs

da.cva

# plots the CVA biplot together with the 95%
# confidence interval circles

plot(da.cva, type = 'n')

## ----------------PCA-------------------- ##
# run PCA on data averaged over judges and replicates

# calculate means using G.R. Hirson's mtable function

mtable<- function (x, bycol, firstvarcol){
      #A function to compute a means table for a
      #matrix.
      #x - the data frame with the data
      #bycol - the row or rows used for grouping
      #(usually wine)
      #firstvarcol - the column containing the first
      #variable

      mns<-matrix(nrow=0, ncol=length(levels(as.
      factor(x[,bycol]))))
```

```
      for (n in firstvarcol:length(x)){
          m.r<-with(x, tapply(x[,n], x[,bycol], mean))
       mns<-rbind(mns,m.r[])
       }
      mns<-as.data.frame(mns)
      names(mns)<-names(m.r)
      rownames(mns)<-names(x[firstvarcol:length(x)])
      mns<-t(mns)
      return(mns)
      }

da.s.m = mtable(da.s, bycol='wine', firstvarcol=4)

# install SensoMineR package from the CRAN first and
# load it

library(SensoMineR)
da.pca = PCA(da.s.m)
```

2.B APPENDIX: R-CODE FOR THE CASE STUDY 2

```
# classical Descriptive Analysis evaluation – example 2
# (c) H. Hopfer, October 2012
# all code comes without any warranty

# reading in your DA data
# 22 judges, 3 replicates, 17 wines (2 AR, 2AU, 1 CH,
# 3 FR, 7PT, 2 US), 18 sensory attributes

da.d = read.table('data2.csv', sep=',', header=TRUE)
head(da.d)
levels(da.d$judge)
levels(da.d$Product)

# define judge, rep and wine as factor
# columns starting with a letter are automatically set
# as factors by R

for(i in 1:3) {
      da.d[,i] <- as.factor(da.d[,i])
      print(is.factor(data[,i]))
      }

dim(da.d) # 22*3*17 = 1122 observations; 18 + 3 = 21
  #columns

# combine all attributes and define it as a matrix

da.a = as.matrix(da.d[,-c(1:3)])
# use all columns but the first three (wine, judge, rep)
head(da.a)
```

```
# building a 3-way MANOVA model with all 2-way
# interactions

da.lm = lm(da.a ~ (judge + Product + rep)^2 , data= da.d)

# run the MANOVA

da.maov = manova(da.lm)

summary(da.maov, test='Wilks')  # print MANOVA table

# in MANOVA wine, rep and judge and the interaction
# wine:judge and rep:judge are significant
# at p< 0.05 => continue with individual ANOVA's

# using the same lm model but now use the ANOVA output

da.aov = aov(da.lm)
aovsum = summary(da.aov)
aovsum # print ANOVA tables for each attribute

# sign. W: FloralA, CitrusA, BananaA, RdFruitA, AlcoA,
# VeggieA, HerbA, SpicyA, VanillaA, CaramelA, ChocoA,
# WoodyA, NuttyA, LeatherA, EarthyA, MedicA, SweetT,
# SourT, BitterT, BodyVisT, AstrinT
# sign. W:J: FloralA, CitrusA, BananaA, RdFruitA,
# VeggieA, HerbA, SpicyA, VanillaA, CaramelA, ChocoA,
# NuttyA, LeatherA, MedicA, SourT,
# sign. W:R: BodyVisT, MedicA, ChocoA, CitrusA

# apply pseudomixed model for attributes with sign.
# wine interactions
# Fnew = (MS_wine)/(MS_wine interaction)

# determine critical F-value for wine effect and W:J
# interaction, df1= df wine and df2 = df W:J

df_W = aovsum[[1]][2,1]  # df for wine effect
df_WJ = aovsum[[1]][4,1]
# df for wine:judge interaction
df_WR = aovsum[[1]][6,1]
# df for wine:rep interaction
newF_crit1 = qf(0.95, df_W, df_WJ)  # critical F value
  #for pseudomixed model of wine and W:J
newF_crit2 = qf(0.95, df_W, df_WR)  # critical F value
  #for pseudomixed model of wine and W:R

aovsum[[1]][2,3]/(aovsum[[1]][4,3]) > newF_crit1
  #floral A still sign.
# ditto for all other attributes with significant wine
# and wine-interaction effects
```

```
# extract sign. attributes only
# sign. W: CitrusA, AlcoA, VeggieA, CaramelA, WoodyA,
# LeatherA, EarthyA, MedicA, SweetT, SourT, BitterT,
# BodyVisT, AstrinT

da.s = da.d[,c(1:4,8:10,12,15:21)]
head(da.s)

# calculate means and LSD

library(agricolae)
cit.LSD = LSD.test(lm(CitrusA ~ (judge + Product +
  rep)^2, data=da.s), 'Product', group=TRUE)
cit.LSD

# ditto for all other significant attributes

# run CVA on the significant data set using the
# candisc package
# install candisc package first from the CRAN and load
# it afterwards

library(candisc)

# build MANOVA model with significant attributes only,
# using only the wine effect
# see Monrozier, R.; Danzart, M. (2001) Food Quality
# and Preference 12:393-406

da.s.mlm = lm(as.matrix(da.s[,-c(1:3)]) ~ Product,
  data = da.s)

da.cva = candisc(da.s.mlm)
da.cva  # extract CVA output => eigenvalues, variance
  #ratios and Bartlett's test for sign. CVs
plot(da.cva, type = 'n')  # plots the CVA biplot
  #together with the 95% confidence interval circles

# consult candisc help for further options regarding
# the plot b
```

REFERENCES

Alasalvar, C., Pelvan, E., Bahar, B., Korel, F., and Olmez, H. 2012. Flavour of natural and roasted Turkish hazelnut varieties (*Corylus avellana* L.) by descriptive sensory analysis, electronic nose and chemometrics. *International Journal of Food Science and Technology* 47: 122–131.

Bacci, L., Camilli, F., Drago, S., Magli, M., Vagnoni, E., Mauro, A., and Predieri, S. 2012. Sensory evaluation and instrumental measurements to determine tactile properties of wool fabrics. *Textile Research Journal* 82: 1430–1441.

Ball, R.D. 1997. Incomplete block designs for the minimisation of order and carry-over effects in sensory analysis. *Food Quality and Preference* 8: 111–118.

Barcenas, P., Elortondo, F.J.P., and Albisu, M. 2000. Selection and screening of a descriptive panel for ewes milk cheese sensory profiling. *Journal of Sensory Studies* 15: 79–99.

Bartlett, M.S. 1947. Multivariate analysis (with discussion). *Journal of the Royal Statistical Society Supplement* 9(B): 176–197.

Beale, E.M.L. and Little, R.J.A. 1975. Missing values in multivariate analysis. *Journal of the Royal Statistical Society B* 37: 129–145.

Cairncross, S.E. and Sjostrom, L.B. 1950. Flavor profiles—A new approach to flavor problems. *Food Technology* 4: 308–311.

Cakir, E., Daubert, C.R., Drake, M.A., Vinyard, C.J., Essick, G., and Foegeding, E.A. 2012. The effect of microstructure on the sensory perception and textural characteristics of whey protein/kappa-carrageenan mixed gels. *Food Hydrocolloids* 26: 33–43.

Chatfield, C. and Collins, A.J. 1980. *Introduction to Multivariate Analysis.* London, U.K.: Chapman & Hall.

Civille, G.V. and Lyon, B.G. 1996. *Aroma and Flavor Lexicon for Sensory Evaluation.* West Conshohocken, PA: American Society for Testing and Materials.

Cliff, M.A. and Heymann, H. 1992. Descriptive analysis of oral pungency. *Journal of Sensory Studies* 7: 279–290.

Dooley, L.M., Adhikari, K., and Chambers, E. 2009. A general lexicon for sensory analysis of texture and appearance of lip products. *Journal of Sensory Studies* 24: 581–600.

Elmaci, Y. and Onogur, T.A. 2012. Mandarin peel aroma: Estimation by using headspace/GC/MS and descriptive analysis techniques. *Acta Alimentaria* 41: 131–139.

Gacula, M.C., Singh, J., Bi, J., and Altan, S. 2008. *Statistical Methods in Food and Consumer Research.* New York: Academic Press.

Garcia-Carpintero, E.G., Sanchez-Palomo, E., Gallego, M.A.G., and Gonzalez-Vinas, M.A. 2012. Characterization of impact odorants and sensory profile of Bobal red wines from Spain's La Mancha region. *Flavour and Fragrance Journal* 27: 60–68.

Heymann, H. 1994. A comparison of descriptive analysis of vanilla by two independently trained panels. *Journal of Sensory Studies* 9: 21–32.

Heymann, H., Machado, B., Torri, L., and Robison, A.L. 2012. How many judges should one use for sensory descriptive analysis? *Journal of Sensory Studies* 12: 111–122.

Heymann, H. and Noble, A.C. 1989. Comparison of canonical variate and principal component analyses of wine descriptive analysis data. *Journal of Food Science* 54: 1355–1358.

Im, J.S., Marshall, R.T., and Heymann, H. 1994. Frozen dessert attribute changes with increased amounts of unsaturated fatty-acids. *Journal of Food Science* 59: 1222–1226.

Keenan, D.F., Brunton, N.P., Mitchell, M., Gormley, R., and Butler, F. 2012. Flavour profiling of fresh and processed fruit smoothies by instrumental and sensory analysis. *Food Research International* 45: 17–25.

King, E.S., Dunn, R.L., and Heymann, H. 2012. The influence of alcohol on the sensory perception of red wines. *American Journal of Enology and Viticulture* 57:481–485.

Lawless, H.T. 1998. Commentary on random vs fixed effects for panelists. *Food Quality and Preference* 9: 163–164.

Lawless, H.T. and Heymann, H. 2010. Chapter 10: Descriptive analysis. *Sensory Evaluation of Foods: Principles and Practices*, 2nd edn. New York: Springer.

Lawless, L.J.R., Hottenstein, A., and Ellingsworth, J. 2012. The McCormick Spice Wheel: A systematic and visual approach to sensory lexicon development. *Journal of Sensory Studies* 27: 37–47.

Lin, S., Hsieh, F., Heymann, H., and Huff, H.E. 1998. Effects of lipids and processing conditions on the sensory characteristics of extruded dry pet food. *Journal of Food Quality* 21: 265–284.

Little, R.J. and Rubin, D.B. 1987. *Statistical Analysis with Missing Data.* New York: John Wiley & Sons.

Lotong, V., Chambers, D.H., Dus, C., Chambers, E., and Civille, G.V. 2002. Matching results of two independent highly trained sensory panels using different descriptive analysis methods. *Journal of Sensory Studies* 17: 429–444.

Lund, C.M., Thompson, M.K., Benkwitz, F., Wohler, M.W., Triggs, C.M., Gardner, R., Heymann, H., and Nicolau, L. 2009. New Zealand Sauvignon blanc distinct flavor characteristics: Sensory, chemical, and consumer aspects. *American Journal of Viticulture and Enology* 60: 1–12.

Machado, B. 2009. Revealing the secret preferences for top-rated dry red wines through sensometrics. MS thesis, University of California at Davis, Davis, CA (Advisor: H. Heymann).

Mammasse, N., Cordelle, S., and Schlich, P. 2011. Do we need replications in sensory profiling? *Rosemarie Pangborn Memorial Meeting*, Toronto, Ontario, Canada. September 4–8.

Mammasse, N. and Schlich, P. 2012. Choosing the right panel size for a descriptive or for a hedonic sensory study based on knowledge gained from two sensory databases. *EuroSense Meeting*, Berne, Switzerland. September 9–12.

Martin, N., Molimard, P., Spinnler, H.E., and Schlich, P. 2000. Comparison of odour profiles performed by two independently trained panels following the same descriptive analysis procedure. *Food Quality and Preference* 11: 487–495.

Meullenet, J.-F., Xiong, R., and Findlay, C.J. 2007. *Multivariate and Probabilistic Analyses*. Ames, IA: IFT Press.

Monrozier, R. and Danzart, M. 2001. A quality measurement for sensory profile analysis: The contribution of extended cross-validation and resampling techniques. *Food Quality and Preference* 12: 393–406.

Munoz, A.M. and Civille, G.V. 1998. Universal, product and attribute specific scaling and the development of common lexicons in descriptive analysis. *Journal of Sensory Studies* 13: 57–75.

Murray, J.M. and Delahunty, C.M. 2000. Selection of standards to reference terms in a Cheddar cheese flavor language. *Journal of Sensory Studies* 15: 179–199.

Nachtsheim, R., Ludi, R., and Schlich, E. 2012. Relationship between product specific and unspecific screening tests and future performance of descriptive panelists. *EuroSense Meeting*, Berne, Switzerland. September 9–12.

Naes, T., Brockhoff, P.B., and Tomic, O. 2010. *Design of Experiments for Sensory and Consumer Data.* New York: Wiley.

Naes, T. and Langsrud, Ø. 1998. Fixed or random assessors in sensory profiling? *Food Quality and Preference* 9: 145–152.

Ng, M., Lawlor, J.B., Chandra, C., Hewson, L., and Hort, J. 2012. Using quantitative descriptive analysis and temporal dominance of sensations analysis as complementary methods for profiling commercial blackcurrant squashes. *Food Quality and Preference* 25: 121–134.

Noronha, R.L., Damasio, M.H., Pivatto, M.M., and Negrillo, B.G. 1995. Development of the attributes and panel screening for texture descriptive analysis of milk gels aided by multivariate statistical procedures. *Food Quality and Preference* 6: 49–54.

Owen, J.G. and Chmielewski, M.A. 1985. On canonical variates analysis and the construction of confidence ellipses in systematic studies. *Systematic Zoology* 34: 366–374.

Parker, M., Osidacz, P., Baldock, G.A., Hayasaka, Y., Black, C.A., Pardon, K.H., Jeffrey, D.W., Geue, J.P., Herderich, M.J., and Francis, I.L. 2012. Contribution of several volatile phenols and their glycoconjugates to smoke-related sensory properties of red wine. *Journal of Agricultural and Food Chemistry* 60: 2629–2637.

Paulsen, M.T., Ueland, O., Nilsen, A.N., Ostrom, A., and Hersleth, M. 2012. Sensory perception of salmon and culinary sauces—An interdisciplinary approach. *Food Quality and Preference* 23: 99–109.

Prindiville, E.A., Marshall, R.T., and Heymann, H. 2000. Effect of milk fat, cocoa butter, and whey protein fat replacers on the sensory properties of low-fat and nonfat chocolate ice cream. *Journal of Dairy Science* 83: 2216–2223.

Robinson, A.L., Adams, D.O., Boss, P.K., Heymann, H., Solomon, P.S., and Trengove, R.D. 2011. The relationship between sensory attributes and wine composition for Australian Cabernet Sauvignon wines. *Australian Journal of Grape and Wine Research* 17: 327–340.

Schlich, P. 1998. Comments from Pascal Schlich on Naes and Langsrud. *Food Quality and Preference* 9: 167–168.

Sjostrom, L.B., Cairncross, S.E., and Caul, J.F. 1957. Methodology of the flavor profile. *Food Technology* 11: A20–A25.

Snitkjaer, P., Risbo, J., Skibsted, L.H., Ebeler, S., Heymann, H., Harmon, K., and Frøst, M.B. 2011. Beef stock reduction with red wine—Effects of preparation method and wine characteristics. *Food Chemistry* 126: 183–196.

Sokolowsky, M. and Fischer, U. 2012. Evaluation of bitterness in white wine applying descriptive analysis, time-intensity analysis, and temporal dominance of sensations analysis. *Analytica Chimica Acta* 732: 46–52.

Steinsholt, K. 1998. Are assessors' levels of a split-plot factor in the analysis of variance of sensory profile experiments. *Food Quality and Preference* 9: 153–156.

Stone, H., Sidel, J., Oliver, S., Woolsey, A., and Singleton, R.C. 1974. Sensory evaluation by quantitative descriptive analysis. *Food Technology* 28: 24–28.

Swaney-Stueve, M. and Heymann, H. 2002. A comparison between paper and computerized ballots and a study of simulated substitution between the two ballots used in descriptive analysis. *Journal of Sensory Studies* 7: 527–537.

Verriele, M., Plaisance, H., Vandenblicke, V., Locoge, N., Jaubert, J.N., and Meunier, G. 2012. Odor evaluation and discrimination of car cabin and its components: Application of the "field of odors" approach in a sensory descriptive analysis. *Journal of Sensory Studies* 27: 102–110.

Wakeling, I.N. and MacFie, H.J.H. 1995. Designing consumer trials balanced for first and higher orders of carry-over effect when only a subset of k samples from t may be tested. *Food Quality and Preference* 6: 299–308.

Warmund, M. and Elmore, J. 2008. A sensory lexicon for black walnut (*Juglans nigra* L.) and Persian walnut (*J. regia* L.). *Hortscience* 43: 1102–1103.

Zeppa, G., Bertolino, M., and Rolle, L. 2012. Quantitative descriptive analysis of Italian polenta produced with different corn cultivars. *Journal of the Science of Food and Agriculture* 92: 412–417.