Chapter 3

Principles of Good Practice

Abstract This chapter outlines the standards of good practice in performing sensory evaluation studies. It briefly covers the sensory testing environment and its requirements, serving samples to panelists, and creating serving procedures, planning. There is a short section on designing experiments including design and treatment structures. Subsequently, it then covers general panelist screening, selecting, and training as well as an overview of panelist incentives. The legal ramifications and requirements of using humans as subjects of sensory tests are also described. Lastly the chapter discusses data collection and tabulation.

Some of the reasons some experimenters advance in trying to resist a (scientific approach) to their work are that: (a). there is no reason to suppose that there will be a bias; (b). it means much more work; (c). things might get mixed up.

There is no reason to suppose that there will not be a bias. As regards (b), one may ask, "more than what?" for that a valid experiment takes more work than an invalid experiment is irrelevant to a man who is wanting to make valid inferences. As regards (c), one feels sympathy, but if an experimenter isn't willing to do a decent job why doesn't he choose some other easier way of earning a living.

—Brownlee (1957, p. 1)

Contents

3.1 Introduction 57
3.2 The Sensory Testing Environment 58
3.2.1 Evaluation Area 59
3.2.2 Climate Control 62
3.3 Test Protocol Considerations 63
3.3.1 Sample Serving Procedures 63
3.3.2 Sample Size 63
3.3.3 Sample Serving Temperatures 64
3.3.4 Serving Containers 64
3.3.5 Carriers 65
3.3.6 Palate Cleansing 65
3.3.7 Swallowing and Expectoration 66
3.3.8 Instructions to Panelists 66
3.3.9 Randomization and Blind Labeling . . . 66
3.4 Experimental Design 66
3.4.1 Designing a Study 66
3.4.2 Design and Treatment Structures 69
3.5 Panelist Considerations 72
3.5.1 Incentives 72
3.5.2 Use of Human Subjects 73
3.5.3 Panelist Recruitment 74
3.5.4 Panelist Selection and Screening 74
3.5.5 Training of Panelists 75
3.5.6 Panelist Performance Assessment 75
3.6 Tabulation and Analysis 75
3.6.1 Data Entry Systems 75
3.7 Conclusion 76
References 76

3.1 Introduction

In later chapters of this textbook we will often state that a particular method should be performed using standard sensory practices. This chapter will describe what we mean by "standard sensory practices." Table 3.1 provides a checklist of many of the good practice guidelines discussed in this chapter; this table can be used by sensory specialists to ensure that the study has

H.T. Lawless, H. Heymann, *Sensory Evaluation of Food*, Food Science Text Series,
DOI 10.1007/978-1-4419-6488-5_3,

Table 3.1 Sensory checklist[a]

Test objective
Test type
Panelist
Recruitment
Method of contact
Supervisory approval
Screening
Informed consent
Incentives
Training
Sample
Size and shape
Volume
Carrier
Serving temperature
Maximum holding time
Test setup
Panelist check-in
Palate cleansers
Instructions
To technicians
To panelists
Score sheets
Instructions
Type of scales
Attribute words
Anchor words
Coding
Randomization/counterbalancing
Booth items
Pencils
Napkins
Spit cups
Clean up
Disposal arrangements (important if security risk)
Receipts if incentive is monetary
Panelist debriefing
Test area
Separation of panelists
Temperature
Humidity
Light conditions
Noise (auditory)
Background odor/clean air handling/positive pressure
Accessibility
Security

[a]This checklist is a quick way of making sure that the sensory specialist has thought of many of the good practice guidelines discussed in this chapter

been thought through. It should be remembered that a good sensory specialist will always follow the standard practices because that would help ensure that he/she will obtain consistent, actionable data. However, an experienced sensory scientist will occasionally break the standard practice guidelines. When one breaks these rules one always has to be fully aware of the consequences, the risks entailed, and whether one still can get valid data from the study.

3.2 The Sensory Testing Environment

Much of the information in this section comes from our experiences in visiting, designing, and operating sensory facilities both in industrial and in university settings. The section was also written with reference to Amerine, Pangborn, and Roessler (1965), Jellinek (1985), Eggert and Zook (2008), Stone and Sidel (2004), and Meilgaard et al. (2006). We feel that anyone planning on constructing or renovating a sensory facility should read Eggert and Zook (2008) and view the accompanying CD, this is an extremely valuable resource. The sensory facility should be located close to potential judges but not in the middle of areas with extraneous odors and/or noise. This means that in a meat-processing plant the sensory area should not be near the smokehouse and in a winery the facility should be out of earshot from the noise of the bottling line. The sensory booth area must be easily accessible to the panelists and if the facility will be used by consumer panelists or panelists that will be traveling some distance then there should be ample, easy parking available. This frequently means that the sensory facility should be on the ground floor of a building and that the area should be near the entrance to the complex. In companies with security concerns, the sensory preparation facility should be within the secure area but the panelist waiting room and possibly the sensory booth area should be in an area that is easily accessible and possibly not secure.

When designing the sensory testing area, the traffic pattern of the panelists should be kept in mind. Panelists should enter and exit the facility without passing through the preparation area or the office areas of the facility. This is to prevent panelists from having physical or visual access to information that may bias their responses. For example, if panelists happen to see some empty jars of a specific brand in a trash can it may bias their responses if they expect to evaluate that brand as one of their coded samples. Additionally, for

security reasons it is not a good idea to have panelists wandering through the sensory area where they may pick up information about projects or other panelists.

3.2.1 Evaluation Area

In its simplest form the facility would need an evaluation area. This may be as simple as a large room that could be used with tables or temporary booths placed on tables. It is always important to remember that if the evaluation occurs in a quiet, uninterrupted manner the likelihood of success is increased. It is especially important that the panelists not influence each other. If temporary booths are not available the sensory specialist should at the very least arrange the tables in room so that the participants do not face each other. Kimmel et al. (1994) arranged a room with tables in such a way that the panelists (in their case children) could not influence each other. If at all possible, separate the panelists with portable plywood booths (see Fig. 3.1 for manufacturing instructions). These can be made inexpensively and will allow panelists to be separated during testing.

Some consumer testing companies use a classroom style where each consumer is seated at a small table with space for a computer screen and the samples. The advantage of this situation is that it is portable (the evaluation area can be set up in hotels, conference rooms, church basements, etc.) and the whole group can receive any verbal instructions simultaneously. If color or appearance is important make sure that the testing area is well lit with balanced daylight-type fluorescent bulbs, however, see Chapter 12 for further information on color evaluation.

In a situation where sensory evaluation is an integral part of the product development and quality assurance cycle of the product a more permanent evaluation area should be constructed. In most sensory facilities, the evaluation area should encompass a discussion area, a booth area and, frequently, a waiting room area for the panelists (Fig. 3.2).

The waiting area should have comfortable seating, be well lit, and clean. This area is often the panelists' first introduction to the facility and should make them feel that the operation is professional and well organized. This area should be modeled on the waiting room of a medical practitioner. The sensory specialist

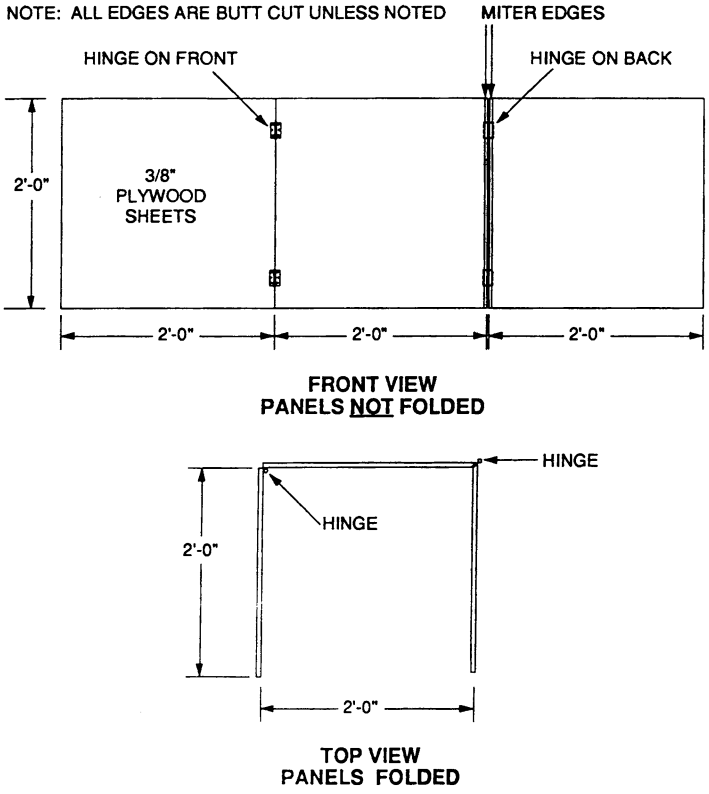

Fig. 3.1 Construction information for portable sensory booths.

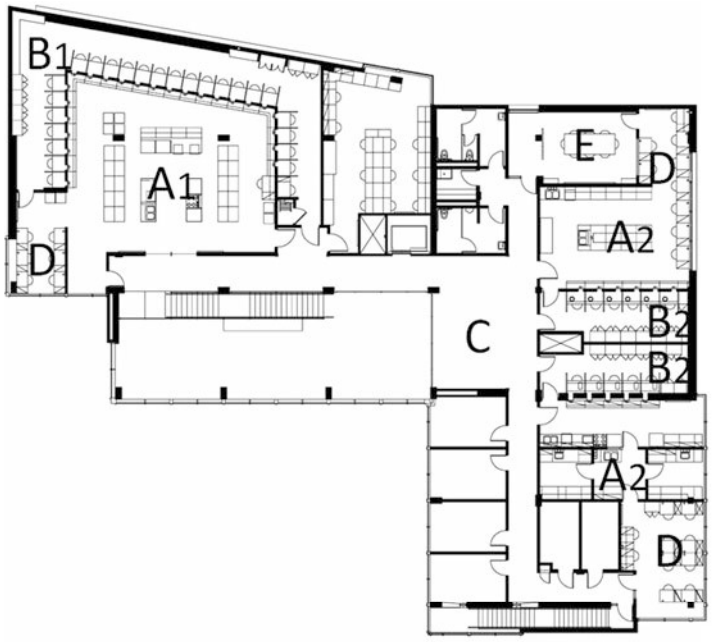

Fig. 3.2 Floor plan of the three sensory facilities in the Robert Mondavi Sensory Building on the University of California at Davis Campus. A1, is a preparation area that includes a range and four ovens; A2 are preparation areas without cooking facilities; B1 is a sensory booth area with 24 individual booths; B2 are two separate sensory booths areas, each with six booths; C is the sensory waiting area with chairs, tables, and sofas; D indicates the work spaces for individual employees and students; E is a focus room with a two-way mirror. (Used with permission from ZGF Architects LLP, Portland, Oregon).

should always try to minimize the waiting time of panelists but sometimes this is unavoidable. To relieve the tedium of waiting the area should be equipped with some light reading. In some facilities, a child care area for panelists' children may also be included. In this case, care must be taken to prevent the noise and distraction from this area from interfering with the panelists' concentration during product evaluation.

In some consumer testing facilities, a briefing area may be adjacent to the waiting room or orientations may be done in the waiting room itself. The orientation area is very useful if chairs in the room are arranged in rows or a semi-circle. Then instructions, as to procedures, can be given to a whole group at once before they enter the test booths or discussion room. Questions can be fielded and volunteer panelists having difficulties can be further instructed or weeded out.

The discussion room would usually be arranged similarly to a conference room but the decor and the furnishings should be simple and in colors that would not affect the panelists' concentration. The area should be easily accessible to the panelists and to the preparation area. However, the panelists should not have visual or physical access to the preparation area. The sections on climate control, lighting, etc., of the booth area are equally applicable to the discussion area.

In many sensory facilities the booth area is the heart of the operation. This area should be isolated from the preparation area, be comfortable but not too casual in appearance. The area should always be clean and professional looking. Once again, neutral or non-distracting colors are advisable. The room should be kept quiet to facilitate panelist concentration. There are probably as many versions of booth areas as there are sensory facilities. Some of the variations are cosmetic and others affect the functionality of the space. In this section we will describe some variations, highlighting advantages and disadvantages of each. We will concentrate on booths used for food evaluations; however, specialized booths are often required for the evaluation of personal care products such as shaving creams, soaps, deodorants and home care products like insecticides, floor waxes, and detergents. An example of a purpose-built sensory booth is Renault motor company's poly-sensorial booth described by Eterradossi et al. (2009).

The number of booths in facilities we have seen ranges from as low as 3 to as high as 25. The number is usually constrained by the space available. However, the sensory scientist should attempt to have the maximum possible number of booths constructed, since the booth availability is frequently a bottleneck in test volume causing undue delay for panelists or decreasing

the number of panelists that can be accommodated. Booth sizes vary greatly in different facilities but the ideal booth is about 1 m by 1 m in size. Smaller booths may make the panelist feel more "cramped" and this could potentially affect concentration. On the other hand, excessively large booths waste space. The booths should be separated from each other by opaque dividers that extend about 50 cm beyond the front edge of the counter top and 1 m above the counter top. This is to prevent panelists in adjacent booths from affecting each other's concentration. The corridor behind the booths should be wide enough for the panelists to comfortably move into and out of the booth area. Additionally, in the United States, if the booths are to be used by disabled persons the guidelines on corridor widths, seating configurations, and counter top heights of the Americans with Disabilities Act of 1990 (42 USC 126 § 12101–12213) should be followed.

The booth counter height is usually either desk or table height (76 cm) or kitchen counter height (92 cm). The height of the booth counter is constrained by the height of the serving counter on the other side of the booth pass-through hatch. We have seen booths where the serving counter was at kitchen counter height and the booth counter was at table height. The potential mess when samples were passed from the higher kitchen counter to the lower booth counter should discourage anyone else from constructing this type of booth. In general, either counter height is used. The table height counters allow the panelists to sit in comfortable chairs but demands that the sensory specialists bend to pass samples through the serving hatches. The amount of bending is minimized when the counter is at kitchen counter height but then the panelists should be supplied stools of adjustable height.

Serving hatches should be large enough to accommodate sample trays, score sheets, and yet small enough to minimize panel observation of the preparation/serving area. The hatches are often about 45 cm wide and 40 cm high; however, the exact size is dependent on the size of the sample serving trays that would be used in the facility. The most popular serving hatches are either the sliding door style or the bread box style. The sliding door style has a door that either slides up or to the side. These doors have the advantage that they do not occupy space either in the booth or on the serving counter. The major disadvantage of these doors is that the panelists can see through the open space into the preparation area. The amount of visual information gleaned by the panelist can be minimized if the sensory specialist stands in front of the open space when serving samples. The bread box design has a metal hatch that is either open to the booth area or to the serving area, but not to both simultaneously. The advantage is that the bread box visually separates the panelist from the serving/preparation area but the disadvantage is that the hatch takes up counter space both in the booth and on the serving counter. The serving hatch should be mounted flush with the counter top, allowing the sensory specialist to easily slide the sample trays into and out of the booth.

The booth should be equipped with electrical outlets for use with computerized data entry systems as well as for electrical appliances that may be needed in the evaluation of a specific product. Data entry systems will be discussed later in this chapter. The installation of sinks in the booth should be discouraged. These sinks are frequently a major source of odor contamination in the booth and are very difficult to maintain in a completely sanitary fashion. It is better to use disposable spittoons and water glasses rather than sinks. When the serving hatch is closed, the panelist should have some means of communicating with the person serving the samples. The ideal communication link is a lighted two-way signal system. In some instances, the communication link can also be an intercom between the booths and the preparation area. In other cases, a card or a simple piece of colored plastic is used, the panelist pushing the signal through a small slot under the serving hatch to gain the attention of the person serving the samples.

Preparation areas will differ based on the product lines evaluated in the particular facility. For example, a facility designed to be used exclusively for frozen desserts would have no need for ovens, but would need ample freezer space. On the other hand, a facility designed for meat evaluations would need freezer and refrigerator space as well as ovens, stove tops, and other appliances used to cook meat. For these reasons it is somewhat difficult to give many rules as to the appliances needed in the preparation area, but there are some appliances and features that would be required in nearly all preparation areas.

The area needs a great deal of storage space. Refrigerated storage is necessary for samples, reference standards, and food treats (incentives) for panelists. Frozen storage space is needed for samples that require freezing. Additionally, cabinet storage space

is required for utensils, serving dishes, serving trays, spittoons, paper ballots, computer printouts of data and statistical analyses, reports, photocopies of literature, etc. Many preparation areas have a lack of adequate storage space. If you as the sensory specialist have any input into the design of a sensory facility insist on ample storage space.

The other area that is often inadequate is the horizontal space required for setting up sensory tests. The counter space should be large enough to allow the specialists to set up one or two sessions' worth of serving dishes at the same time. The space can be re-used if food service trays and vertical food service carts are used as a holding space prior to serving samples. The entire area should be constructed with materials that are easy to clean and to maintain. Dishwashers, sinks with garbage disposals and trash cans should be installed in the preparation area. There should also be an adequate clean water supply for cleaning purposes as well as a supply of tasteless and odorless water to be used by the panelists for rinsing between samples. Double distilled water or bottled water from a reputable dealer are generally preferred. Additionally, depending on the types of products to be tested, other appliances such as electrical or gas cook tops and ovens, microwave ovens, deep-fat fryers may also be required. If oven and cook tops are installed, the area then requires hoods with charcoal filters or outside venting to control odors from the cooking area(s). The list of possible appliances is almost endless. In some facilities, flexibility has been designed into the preparation area with moveable case goods, flexible electrical and water hookups, and the potential to roll in new appliances and remove appliances that are not needed for a specific test. Again, space for storage of specialized equipment such as rice cookers or tea pots must be considered when designing a facility. In addition, local restaurant building codes should be consulted to make sure that sprinklers (used for fire safety), water quality, sewer, and all other utilities are adequate in the preparation area.

3.2.2 Climate Control

The booth and discussion areas should be climate controlled and odor free. The use of replaceable active carbon filters in the ventilation system ducts supplying these areas is encouraged. These areas should have excellent ventilation. A slight positive air pressure in these areas can minimize odor transfer from the preparation area. The sensory scientist should make sure that any cleaning supplies used in the booth and discussion areas do not add extraneous odors. These areas should be as noise free and distraction free as possible. Signs requiring silence in the hallways around these areas during testing times are helpful. Additionally, the noise added by nearby mechanical systems, e.g., freezers, air conditioners, processing equipment should be minimized.

The temperature and relative humidity for the booth and discussion areas should be 20–22°C and 50–55% relative humidity. These conditions would make the environment comfortable for the panelists and will prevent them from being distracted by the temperature or the humidity.

Illumination in these areas should be at least 300–500 lx at the table surface. Ideally it should be controllable with a dimmer switch to a maximum of 700–800 lx, the usual illumination intensity in offices. Incandescent lights are modifiable, by changing light bulbs, and versatile by allowing one to control both the light intensity and the light color. However, heat buildup can be a problem and should be accounted for when designing the booth area. The lighting should be even and shadow free on the counter surface. There are special lighting requirements for color evaluations and these will be discussed in Chapter 12.

The above discussion was for an average sensory facility used for food testing. However, for some product ranges more specialized facilities should be constructed. For example, if the facility is to be used to test ambient odor, for odor thresholds, room air deodorants, odors associated with household cleaners, etc., then either an odor room or a dynamic olfactory testing area should be created.

A dynamic olfactory test area would contain an olfactometer. In an olfactometer a gaseous sample flows continuously through tubes and the sample is diluted by mixing with odor-free air. Panelists would evaluate the samples at an exit port using a face mask or specially designed sniffing ports (Takagi, 1989). An odor room can be used by more than one panelist simultaneously. The odor evaluation area consists of an anteroom and a test room. The anteroom shields the test chamber from the external environment. The odor area should be constructed of odor free,

easy to clean, non-absorptive materials. Stainless steel, porcelain, glass, or epoxy paints may be appropriate. The test room should have a ventilation system that can completely remove odorized air and introduce a controllable odor-free background.

3.3 Test Protocol Considerations

3.3.1 Sample Serving Procedures

The sensory specialist should be very careful to standardize all serving procedures and sample preparation techniques except the variable(s) under evaluation. For example, in a study to evaluate the effect of accelerated ripening on cheddar cheese flavor we decided to do a triangle test. Two technicians were assigned to the project and they divided cutting the cheese samples into 1 cm^3 cubes, each technician cutting either cheese A or cheese B. One technician was very precise and all the cubes she cut were exactly 1 cm^3. The other technician was less precise and her cubes varied slightly in size. Once the cubes were placed in serving containers it became obvious that the panelists could identify the odd sample by visual inspection only. These cheese samples were thrown away and the more precise technician was assigned to do all the cubing. However, she could not cut the cheese and serve it immediately. The cubes had to be stored overnight in a refrigerator. The technician decided to store the cubes from cheese A in one refrigerator and those from cheese B in another. The two refrigerators varied slightly in their temperature settings. The next morning, the samples were served to the panelists, who by simply touching the samples could identify the odd sample. The samples then had to be stored again, but in the same refrigerator to equilibrate the temperature difference. In a different study, involving threshold determination by discrimination testing, the layout of the samples prior to serving along a benchtop allowed a temperature decrease in samples set nearer to a room air conditioner. Some panelists could pick out the samples that were different based on this small temperature difference.

If carriers or combinations of products are required the timing of this process must be standardized. For example, if milk is poured on a breakfast cereal, the amount of time between pouring and tasting must be the same for all samples. It may not be wise to simply pass the milk in a container into the test booth for the panelists to add without instructions. They may pour the milk on all the samples at the outset, with the result that the last one evaluated has a much different texture than the first.

As can be seen from these examples, the sensory specialist should pay careful attention to the following areas when writing the test protocol and when performing the study: the visual appearance of the sample, sample size and shape, and sample serving temperature. Additionally, the sensory specialist should decide which serving containers should be used, whether the sample should be served with a carrier, how many samples should be served in a session, whether the panelists should rinse their mouths between samples, whether samples are to be expectorated or swallowed and how many samples should be served in a session. In the following sections each of these issues will be discussed, many of the suggestions made in these sections are based on our own experiences in a variety of sensory settings.

3.3.2 Sample Size

If the samples are evaluated in a discrimination test and the appearance of the sample is not the variable under evaluation then the samples should appear identical. If it is not possible to standardize the appearance exactly, a sequential monadic serving order may be used (Stone and Sidel, 2004). However, if there is a possibility that the panelists may remember that the samples were not identical in appearance then a discrimination test is not appropriate.

Cardello and Segars (1989) found that sample size affected the intensity scores assigned to textural attributes by panelists, even when the panelists were unaware of the sample size differential. De Wijk et al. (2003) confirmed these results with a different product. These results make it very important that the sensory specialist specifies the sample size and shape used in study, since the possibility exists that a different sample size may have led to different results. Therefore, when deciding on the sample size to serve the sensory specialist should keep a few questions in mind, namely What is the purpose of this study? How large is the normal portion size for this product? How large is a

normal mouthful of this product? How many attributes does the panelist have to evaluate on this product? Is it possible to easily manipulate the size of the product? The answers to these questions should lead the sensory specialist to a reasoned decision in determining the size of the sample to be served. Keep in mind that it is better to err slightly on the side of a more generous portion size than a stingier one. In some cases a minimum amount to be eaten may be specified. This is potentially important in consumer tests where some participants may be timid about tasting novel products. However, a reasonable balance between cost associated with the product, storage, and preparation in relation to the sample size should be maintained.

3.3.3 Sample Serving Temperatures

The serving temperature of the product must be specified in the test protocol. Serving temperatures and holding time can present difficulties with some products such as meats. One approach to this is to serve the items in containers that are themselves warmed. In our laboratories and others, sand baths heated in an oven to a fixed temperature (usually 50°C) are used. Small glass beakers or ceramic crucibles used as holding dishes sit embedded in the sand baths and these in turn hold the samples to be tested. Even with this arrangement it is important to minimize the time samples are held or at the very least maintain this time as a constant across panelists.

In dairy products such as fluid milk, sensory characteristics may be accentuated if the product is warmed to a temperature above those of storage. In some tests where sensitivity and discrimination are the primary concerns, this is less realistic but a serving temperature allowing better discrimination is warranted. Thus fluid milk can be served at 15°C instead of the more usual 4°C to enhance the perception of volatile flavors. Ice cream should be tempered at –15°C to –13°C for at least 12 h before serving since scooping is difficult if the ice cream is colder. At higher temperatures the ice cream would melt. It is also usually best to scoop ice cream directly from the freezer immediately before serving rather than to scoop the portions and store these in a freezer. In this latter case the surface of the sample portion is inclined to become icier than the outer portion of a freshly scooped sample.

When samples are served at ambient temperatures the sensory specialist should measure and record the ambient temperature during each session. For samples served at non-ambient the serving temperature should be specified as well as the method of maintaining that temperature, whether it is sand baths, thermos flasks, water baths, warming tables, refrigerators, freezers, etc. The temperatures of samples that are served at non-ambient temperatures should be checked at the time of serving to ensure that the specified temperatures were achieved. Additionally, the specialist should specify the sample holding time at the specified temperature.

If samples are to be held for an extended period, the test protocol should include a discrimination test, with sufficient power (see Chapter 4) to determine if the holding period leads to changes in the sensory attributes of the product. If no changes occur then the samples could be held for an extended period. However, if products are to be held at elevated temperatures for any period the sensory specialist should also monitor potential microbial growth that could compromise the safety of the panelists.

3.3.4 Serving Containers

It is difficult to give rigid rules as to the choice of containers since different conditions exist in different sensory facilities. In some facilities, it is expensive and time consuming to wash many dishes, thus specialists in these cases would tend toward using disposable containers. In other facilities, there may be financial or environmental constraints that preclude the use of disposable dishes. The best advice is to use common sense when deciding which containers to use. The sensory specialist should choose the container that is most convenient, yet the choice of container should not negatively affect the sensory attributes of the product. For example, Styrofoam cups are very convenient to use since they are disposable and can easily be labeled using either a permanent ink marker or a stick-on label, yet we have found that these containers can adversely affect the flavor characteristics of hot beverages. If three-digit codes are applied via marking pens, care must be taken to insure that the ink does not impart an aroma.

3.3.5 Carriers

The issue of whether or not to use carriers poses some problems for the sensory specialist and deserves careful consideration. "Carriers" usually refer to materials that form a base or vehicle for the food being tested, but may more broadly be considered as any other food that accompanies the one being tested so that they are ingested (and tasted) together. Examples are cream fillings in pastries, butter on bread, spices in a sauce, and salad dressing on lettuce leaves.

In discrimination testing, the goal is often to make a test that will be very sensitive to product differences. A carrier can mask or disguise differences or minimize panelists' abilities to perceive the difference due to the addition of other flavors and modifications to the texture and mouthfeel characteristics. In some cases the carrier may simply increase the overall complexity of the sensory impressions to the point where the panelists are overwhelmed. In these cases the use of a carrier might not be desirable since it will decrease the effective sensitivity of the test for detecting sensory differences. If there are serious consequences from missing a difference (Type II error, see Chapter 4) then the use of a carrier that could potentially mask differences is not recommended.

If on the other hand, a false alarm or detection of a false positive difference (Type I error, see Chapter 4) poses serious problems, then the obscuring of a difference by the carrier is less detrimental. The degree of realism added by the carrier may complicate the situation, but it could prevent the detection of a difference that might be meaningless to consumers. The sensory specialist should discuss with the client whether the degree of realism in the test is a concern. For a food product that is rarely consumed alone and almost always involves a carrier, the "artificiality" of the situation where the carrier is omitted may be a major psychological problem to the panelists, especially in consumer testing. An example would be cherry pie fillings which are rarely eaten without pie crust.

So, there are two guidelines for consideration in determining whether a carrier should be used: the relative consequences of missing a difference versus a false positive test result and the degree of realism that is deemed necessary. Often the complications created by use of the carrier will lower the degree of sample control and uniformity that is possible, so this must be considered. Careful discussion of these issues with the client can help clarify the best approach. In some cases it may be advisable to do the test both with and without the carrier if time and resources permit. This can be very informative about the size of the perceivable difference as well as the nature of the interactions between the carrier and the food to be tested.

Stone and Sidel (2004) give the following interesting example of a compromise in the use of a carrier, where the food product (a pizza sauce) is influenced by the carrier (crust) in such important ways that the preparation, but not the testing, had to involve the substrate: "... it was agreed that flavor interactions with crust components resulted in a sauce flavor that could not be achieved by heating the sauce alone. However, owing to variability in pizza crust within a brand, it was determined that the pizza would be cooked and the sauce scraped from the crust and tested by itself. Thus the chemical reactions were allowed to occur and the subjects' responses were not influenced by crust thickness or other non test variables."

3.3.6 Palate Cleansing

The goal of palate cleansers should be to aid in the removal of residual materials from previous samples. An anecdote frequently told at wine tasting events says that serving rare roast beef slices will help undo the effects of high tannin in red wine samples. This makes some sense chemically. The proteins of the meat and its juices could form a complex removing tannins from solution—reducing the "pucker" of the wine. There have been numerous studies on palate cleansers to remove red wine astringency (see Ross et al., 2007 for a relatively recent example). However, it would seem that no true consensus has been reached on the ideal palate cleanser to use in these conditions.

Lucak and Delwiche (2009) evaluated the effects of a range of palate cleansers (chocolate, pectin solution, table water crackers, warm water, water, and whole milk) on foods representing various tastes and mouthfeel effects such as jelly beans (sweet), coffee (bitter), smoked sausage (fatty), tea (astringent), spicy tortilla chip (pungent), mint (cooling), and applesauce (non-lingering). They found that table water crackers were the only palate cleanser effective across all representative foods.

A study of off-flavors in fish examined the difficulties panelists have when cleaning the palate of methyl isoborneol, a compound associated with earthy, muddy, or musty aroma (Bett and Johnson, 1996). They suggested the use of untainted fish itself as the cleanser to use between test samples. This would make sense in that the fish flesh is an effective binder of the odor compound in question. However, these authors did raise the concern that this would involve time and expense in using additional fish samples as a palate cleanser.

3.3.7 Swallowing and Expectoration

In most analytical sensory tests, swallowing is avoided and samples are expectorated. This is assumed to provide less carry-over or unwanted influence of one product to the next. Also swallowing high-fat products can add unnecessary calories to panelists' diets. Of course, in consumer testing where acceptability is being measured, swallowing and post-ingestion effects can affect consumers' opinions on the products. Also in consumer testing generalizing to the natural consumption is a concern and here having respondents swallowing the products is acceptable. Kelly and Heymann (1989) studied the effect of swallowing versus expectoration on thresholds and fatigue effects in paired comparison and triangle tests using added salt in kidney beans and added milk fat to skim milk. They found no significant effects. However, it should be noted that the power of the test was low and thus the likelihood of finding a difference was slight. A time–intensity evaluation of Yerba mate infusions by Calviño et al. (2004) found that swallowing versus expectoration did not affect the perceived bitterness intensity of the infusion but that spitting did increase the rate of decay of the sensation.

One advantage of swallowing in analytical sensory testing is the stimulation of sensory receptors in the throat. This can be important in some products and flavor systems. For example, throat burn is important in pepper samples and "throat catch" (another type of chemical irritation) is characteristic of chocolate.

3.3.8 Instructions to Panelists

These should be very clear and concise. It is frequently desirable to give the instructions on how to perform the sensory evaluation both verbally, before the panelists enter the booth area, and in written form on the score sheet. These instructions should be pre-tested by having someone unfamiliar with sensory testing and the project attempts to follow them. We have frequently been amazed at how easily panelists misread or misunderstand what seemed to us to be simple, clear instructions. This usually occurs because we are too familiar with the testing methodology and thus read more into the instructions than is really there. The sensory specialist should always be aware of this potential problem.

The instructions to technicians and staff should also be very clear and preferably should be written. It is useful to have the technicians repeat the explanation of the procedure to the sensory specialist. This will assure that there were no communication gaps between the sensory specialist and the people performing the study. Additionally, for many tests it is useful to develop a standard operating procedure and to keep this available in laboratory notebooks.

3.3.9 Randomization and Blind Labeling

Samples should be blind labeled with random three-digit codes to avoid bias and sample order should be randomized to avoid artifacts due to order of presentation. Table 3.2 gives step-by-step instructions to set up discrimination and preference tests and Table 3.3 does the same for rating, ranking, and hedonic tests. Figures 3.3, 3.4, 3.5, 3.6, and 3.7 show master sheets prepared according to the instructions in Tables 3.2 and 3.3.

3.4 Experimental Design

This chapter is not designed to be a comprehensive discussion on experimental design. Excellent books and chapters have been written on experimental design and we would refer the reader to Cochran and Cox (1957), Gacula and Singh (1984), Milliken and Johnson (1984), MacFie (1986), Petersen (1985), Hunter (1996), and Gacula (1997).

3.4.1 Designing a Study

In this section we want to highlight some major issues that should be kept in mind by the sensory specialist when designing an experiment. At the beginning

Table 3.2 Step-by-step instructions for setting up discrimination and preference tests

1. Prepare master sheet (see the completed master sheets in Figs. 3.3, 3.4 3.5, and 3.6).
   a. Fill in the sample identification at top. For paired difference or paired preference tests two columns should be filled in (A, B). For constant reference duo-trio tests three columns should be indicated (Reference A, A, and B). For balanced reference duo-trio tests four columns are needed (Reference A, Reference B, A, and B). Triangle tests also need four columns filled in (A, A, B, and B). Only the researchers should know the identity of the A and B.
   b. Fill in judge numbers (i.e., 1, 2, 3. . .). Assign each judge a number and make sure that a key to these numbers is included in the notebook associated with study. It is simpler if a specific judge retains that number throughout the study.
   c. Create permutations of sample presentation. For paired difference or paired preference tests there are two possible permutations (AB, BA). For the balanced reference duo-trio tests there are four possible permutations (R_A AB, R_A BA, R_B AB, R_B BA).and the constant reference duo-trio tests has two possible permutations (R_A AB, R_A BA). For triangle tests there are six possible permutations (AAB, ABA, BAA, BBA, BAB, ABB). Each serving order should be assigned a number.
   d. Determine order of sample presentation. Using a table of random permutations, numbers are read from top to bottom within a column. Use only numbers corresponding to the number of serving orders in the test. Write the number (with a red pen, indicated in **bold** on Figs. 3.3, 3.4, 3.5, 3.6 and 3.7) in a blank column and then write the order that the samples will appear on the serving tray in the upper right hand corner of each square on the master sheet. This indicates the order in which each sample is presented to each judge.
   e. Assign three-digit random code numbers to each sample for each judge. Start from any point on the table of random numbers and use three digits for each number. Do not use numbers that may have meaning to the judges (i.e., 13, 666, 911). Write the random numbers on the master sheet, one for each sample for each judge (use blue or black pen, indicated in *italics* in Figs. 3.3, 3.4, 3.5, 3.6 and 3.7). An occasional duplicate of a number may be found on a random number table, if so, skip the duplicate number.
2. Write random number codes on the sample containers. Use the random code numbers which were written on the master sheet. Code numbers on sample containers should match the appropriate code numbers on the master sheet. The sample containers to be filled with the reference samples should not be coded R_A or R_B but should be coded only with an R.
3. Prepare score sheet. Fill in the date, the judge number, and the random code numbers in the sequence in which the samples are to be evaluated (as indicated by the random permutations).
4. Prepare samples.
   a. Prepare an organized arrangement for portioning samples. A simple method is to make a master sheet template with sufficient space for the sample containers to be placed in the squares. This template may be made out of any large paper or available substitute. Allowing a 3 in. square for each sample is suggested, however, this will vary depending on the sample container itself.
   b. Assemble sample containers on template. Once all the containers are placed on the template it should be identical in appearance to the master sheet.
   c. Portion samples into containers.
5. Assemble samples for each judge on a tray in the sequence that they are to be evaluated. Also, place the score sheet on the tray and water for rinsing the palate. Double check serving order.
6. Serve samples to judges for evaluations.
7. Decode score sheet on the master sheet. Circle the code that the judge circled. (Use a pen, *never* use pencil on master sheets or score sheets). In this way, decoding is simple and orderly. In order to analyze the data, it must be represented numerically. This may be according to the number of correct judgments (paired difference test, triangle test, and duo-trio test) or number of judges preferring sample A or B (paired preference). Make sure that a column is left for this purpose.
8. Analyze the data.

of any project the sensory specialist and all the parties that are requesting the study should define the objective of the study. To ensure that all parties are clearly communicating, the sensory specialist should rephrase all the objectives as questions. These should be circulated among all parties, who should provide feedback to the sensory specialist. The sensory specialist, in consultation with the client(s), should identify the tests required to answer the questions. At this point it is usually instructive for the specialist to design the perfect experiment without any cost constraints. This exercise is instructive because the process allows the specialist to clearly indicate the "ideal." Then when time and cost constraints are added and the specialist has to redesign the study to a scaled down version there is a clear picture of what is "given up" in this process. In some situations the scaled down version may not be capable of answering the test objectives. When this happens, the specialist and the client(s) must renegotiate the cost and time constraints and/or the test objectives. It is usually better to decrease the number of test objectives to those with the highest

Table 3.3 Step-by-step instructions for setting up ranking, rating and hedonic tests

1. Prepare master sheet (see Fig. 3.7 for completed master sheet).
   a. Fill in sample identification at top. In the example, in Fig. 3.7 for a study of fish, this may be Scrod, Cod, Tuna, Hake. Only the researchers should know the identity of the products or samples.
   b. Fill in judge numbers (i.e., 1, 2, 3. . .). Assign each judge a number and make sure that a key to these numbers is placed in the study notebook. It is simpler if a specific judge retains that number throughout the study.
   c. Assign three-digit random code numbers to each sample for each judge. Start from any point on the table of random numbers and use three digits for each number. Never use numbers that may have meaning to the judges (i.e., 13, 666, 911). Write the random numbers on the master sheet, one for each sample for each judge (use blue or black pen, indicated in *italics* in Fig. 3.8). An occasional duplicate of a number may be found on a random number table. If so, skip the duplicate number.
   d. Determine order of sample presentation. Using a table of random permutations, numbers are read from top to bottom within a column. Use only numbers corresponding to the number of samples being tested (i.e., for four samples: use only numbers 1, 2, 3, and 4; read the numbers in the order they appear). Write the number (with a red pen, indicated in bold in Fig. 3.7) in the upper right-hand corner of each square on the master sheet. This indicates the order in which each sample is presented to each judge. In the example, the first sample is served fourth, the second sample is served first, etc. for judge 1.
2. Write the random codes on the sample containers. Use the random code numbers which are written on the master sheet. Code numbers on sample containers should match the appropriate code number on the master sheet. If there are enough people working together, this can be done as random numbers are recorded on the master sheet.
3. Prepare score sheet. Fill in the date, the judge number, and the random code numbers in the sequence in which the samples are to be evaluated (as indicated by random permutations).
4. Prepare samples.
   a. Prepare an organized arrangement for portioning samples. A simple method is to make a master sheet template with sufficient space for the sample containers to be placed in the squares. This template may be made out of any large paper or available substitute. Allowing a 3 in. square for each sample is suggested, however, this will vary depending on the sample container itself.
   b. Assemble sample containers on template. Once all the containers are placed on the template it should be identical to the master sheet.
   c. Portion samples into containers.
5. Assemble samples for each judge on a tray in the sequence that they are to be evaluated. Also, place the score sheet on the tray and water for rinsing the palate. Double check serving order.
6. Serve samples to judges for evaluations.
7. Decode score sheet on the master sheet. When judges are asked to rate only one attribute a blank column is left between columns of random code numbers. When asked to rate more than one term more blank columns (one column for each terms rated) should be left. These columns provide space for recording judge scores after completion of the test. (Use a pen, *never* use pencil on master sheets or score sheets). In this way, decoding is simple and orderly.
8. Analyze the data.

priorities rather than cutting the power of the test (see Appendix E for power issues). There is no point in performing a study that is inadequate in answering the major test objectives. If it is not possible to design an adequate study, the specialist must ask for more resources.

Next, the sensory specialist should meticulously scrutinize the study step by step. The idea is to ask questions at each point about the worst possible scenario and how the study could be improved to minimize these contingencies. Sensory studies are more complex than they appear at first glance and the potential for complications and mistakes is always present. Samples may be lost, contaminated, or otherwise mishandled. Panelists may drop out before completing the test sequence. Participants may not correctly follow the test protocol or they may misunderstand instructions. Technical personnel can make mistakes in serving order sequences. Unwanted fluctuations in sample temperature or other conditions may enter the picture. Most of these problems can be eliminated or minimized in a well-designed test.

Once the study has been redesigned it is a good idea to write down a "skeleton" statistical analysis. This will give the specialist a good idea about the degrees of freedom associated with significance tests. It is also helpful to sketch out potential figures and tables that will be used in the final report.

Permutation numbers (Perm #) AB = 1
BA = 2

Judge	Perm. #	A	B		
1	2	169^{2}	507^{1}		
2	1	212^{1}	194^{2}		
3	1	962^{1}	644^{2}		
4	2	273^{2}	693^{1}		
Etc.					

Fig. 3.3 Example of a master sheet for a paired preference test.

Permutation numbers (Perm #) R_A AB = 1
R_A BA = 2

Judge	Perm. #	R_A	A	B	
1	1	*R*	557^{1}	485^{2}	
2	2	*R*	636^{2}	684^{1}	
3	1	*R*	325^{1}	238^{2}	
4	2	*R*	401^{2}	159^{1}	
etc..					

Fig. 3.4 Example of a master sheet for a constant reference duo–trio test.

3.4.2 Design and Treatment Structures

We like to use description of the experimental design elucidated by Milliken and Johnson (1984). These authors divided experimental design into two basic structures, namely treatment structure and design structure. They describe the treatment structure as set of samples or treatments that the client(s) selected to study in the specific project. The design structure

Permutation numbers (Perm #) R_A AB = 1
R_A BA = 2
R_B AB = 3
R_B BA = 4

Judge	Perm. #	R_A	R_B	A	B
1	4		R^{1}	557^{3}	485^{2}
2	1	R^{1}		636^{2}	684^{3}
3	2	R^{1}		325^{3}	238^{2}
4	3		R^{1}	401^{2}	159^{3}
etc..					

Fig. 3.5 Example of a master sheet for a balanced reference duo–trio test.

Permutation numbers (Perm #) BAA = 1 BBA = 4
ABA = 2 BAB = 5
AAB = 3 ABB = 6

Judge	Perm. #	A	A	B	B
1	5		495^{2}	926^{1}	183^{3}
2	4	292^{3}		899^{1}	854^{2}
3	2	797^{1}	630^{3}	315^{2}	
4	3	888^{1}	566^{2}	981^{3}	
5	1	267^{2}	531^{3}	469^{1}	
6	6		201^{1}	239^{2}	827^{3}
etc..					

Fig. 3.6 Example of a master sheet for a triangle test.

is defined by sensory specialists when they group experimental units into blocks. These two structures are linked by the randomization performed by the

Judge	Scrod	Cod	Tuna	Hake	
1	*909*[4]	*623*[3]	*703*[2]	*903*[1]	
2	*690*[1]	*558*[2]	*578*[3]	*383*[4]	
3	*694*[3]	*373*[1]	*693*[4]	*290*[2]	
4	*890*[2]	*763*[4]	787[1]	661[3]	
etc..					

Fig. 3.7 Example of a master sheet for a rating, ranking of hedonic test.

sensory specialist prior to the study and together they make up the experimental design of the study. The sensory specialist should let treatment structure dictate neither a poor design structure nor a favorite or frequently used design structure affect the selection of treatments.

3.4.2.1 Design Structures

Completely Randomized Design (CRD)

In this design all the samples are randomly assigned to all the panelists. Most of the experimental designs associated with sensory studies are performed to avoid or minimize artifacts due to order of sample presentation. The simplest solution to this problem is to make sure that the sample presentation order is completely randomized across all panelists. This technique works quite well in situations where the number of samples is small and all samples can be evaluated by all panelists in a single session. CRD is the ideal design for a central location consumer test where each panelist evaluates each sample. For example, in a mall intercept, test panelists are asked to express their degree of liking for each of four cola products. Each panelist receives the four colas in a randomly assigned sequence.

CRD designs also include random assignment of products to people where each individual only sees one product. These so-called consumer monadic tests are common in consumer field studies. These are also called between-groups comparisons since there are different groups of people evaluating each product. The product group forms a block. An example would be a study with three versions of a product. The total consumer group is divided into three subgroups with people randomly assigned to a group. Each group tests one product, then fills out a questionnaire. Justification for monadic designs arises when (1) the test would be too time consuming or lengthy to have all people evaluate all products; (2) the use of one product would be likely to influence opinion of another; or (3) the use of the product changes the environment, person, or substrate. The last effect is common with consumer products (e.g., floor wax, insecticide) and personal care products (e.g., skin cream, hair conditioner). Time pressure to complete a test might also dictate a monadic design in consumer field work.

With trained analytical panels, the samples should be evaluated in replicate (often triplicate) to ascertain judge to judge variation. If the number of samples is sufficiently small it may be possible to have each panelist evaluate all samples in replicate in a single session using CRD. However, this is often not possible and then the sensory scientist would use a randomized complete block design (RCBD).

Randomized Complete Block Design (RCBD)

In a randomized complete block design each treatment (usually samples) is randomly assigned to each unit (usually panelists) within each block (often sessions). This design is frequently used when trained analytical panelists cannot evaluate all samples in replicate in a single session. In this case the best solution is to have each panelist evaluate all samples in a single session and then have them return for a subsequent session to re-evaluate all the samples. An example is a descriptive analysis study of six ice creams made with fat replacers. In a single session the panelists can only evaluate six samples. However, the samples should be evaluated in triplicate. The panelists must attend three sessions to evaluate all the samples in triplicate. In this study the blocks are the sessions and the six samples are randomized across those panelists within each block.

Incomplete Block Design

Incomplete block designs are used when there are too many treatments in the experiment for the panelists to

judge all samples in a single session (block). In this case the panelists evaluate subsets of samples in individual sessions. The objective may be for each panelist to ultimately evaluate all samples often in replicate or it may be that panelists only see a subset of samples. An example of the first type of incomplete block design is the descriptive analysis of 13 vanilla samples performed by one of the authors (Heymann, 1994). The panelists could not evaluate all 13 samples in a single session. We chose an incomplete block design with four samples per block (session) and 13 blocks (sessions) (plan 11.22, Cochran and Cox, 1957). At the end of the study all the panelists had evaluated each of the 13 vanillas four times. The second type of incomplete block design is often used in consumer studies where the purpose is to screen flavor or fragrance candidates from a large pool of potential flavors or fragrances. For example, there may be twenty eight possible fragrances for a new floor wax, but due to the fatiguing nature of the fragrances consumers cannot rate their liking for more than four fragrances in a session. By choosing the appropriate incomplete block design (plan 11.38, Cochran and Cox, 1957) 63 groups of nine consumers would evaluate four fragrances in a screening test to pick the most liked fragrances.

3.4.2.2 Treatment Structures

One-Way Treatment Structure

In this case a set of treatments (samples) are chosen without assuming a relationship among the treatments. In sensory studies this occurs when a product set is chosen from among the brands on the market. In these cases there is no assumption the product made by one company is related to that from another company. For example, in a study of the sensory characteristics of black tea, the sensory specialist may choose four black teas, one made by each of four nationally known companies. The one-way treatment structure for this study then has four samples that are not related to each other in any way except that they are national brands of tea.

Two-Way Treatment Structure

For two-way treatment structures a set of samples are created by combining levels of two different types of treatments. In the sensory setting one may choose a product set from among the brands on the market and then each of these products is prepared in two different ways prior to sensory evaluation. For example, to return to the tea example used above, the sensory specialist decides to evaluate the teas as a hot beverage and as an iced tea. The treatment structure for this study is then two way with a total of eight treatments (four teas at two temperatures). Two-way treatment structures are also known as factorial treatment structures.

Other Treatment Structures

Many other possible treatment structures exist and are used in sensory studies. Examples include fractional factorial structures, a one-way structure of controls combined with a two-way factorial arrangement treatment structure. Split-plot and repeated measures experimental designs are created from incomplete block design structures and factorial arrangement treatment structures with two or more types of treatments. In a simple split-plot experimental design there are two sizes of experimental units and the treatments can be applied to differently sized experimental units by randomization. An example would be the following: each of six varieties of potatoes is grown in three rows, randomly assigned, in a field; the potatoes are harvested and the potatoes from each row are kept in separate containers. The potatoes are the cooked, using three cooking techniques. Each container is split into three batches and randomly assigned a cooking procedure. The cooked potatoes are evaluated by a descriptive analysis panel for texture. In this case the experimental unit to evaluate variety is the row and the experimental unit to evaluate the cooking procedure is the batch (Milliken and Johnson (2004).

A simple repeated measures designs is similar to a simple split-plot designs in terms of the two sizes of experimental units but the levels of at least one treatment (usually time) cannot be randomly assigned. For example, broccoli is harvested and randomly assigned to be packaged in four different packaging materials. The packages are stored and a sensory descriptive panel evaluates the samples daily in triplicate over a 2-week period. In this case the one experimental unit is the packaging type and the other is time (daily).

The following are some of the textbooks with numerous examples of more complex treatment structures: Milliken and Johnson (1984), Cochran and Cox (1957), Petersen (1985).

3.4.2.3 Randomization

The setting up instructions in Tables 3.2 and 3.3 only indicate how sample order may be randomized. It is usually better to also ensure that sample order is counterbalanced, as far as possible. When sample order is counterbalanced each serving sequence occurs an equal number of times. To determine if a specific master sheet is counterbalanced, one must determine the number of times each serving sequence appears. In a fully counterbalanced design all potential serving sequences will occur an equal number of times. It is possible to use specially designed serving sequences allowing the sensory scientist to not only have completely counterbalanced designs but also have serving sequences that are completely balanced. In other words, every sample is preceded by every other sample an equal number of times (MacFie et al., 1989; Wakeling and MacFie, 1995). These designs are especially helpful when the possibility of carry-over effects between samples exist (Muir and Hunter, 1991/1992; Schlich 1993; Williams and Arnold, 1991/1992). They are also helpful as "insurance" against carry-over effects, since their use allows one to determine carry-over effects post hoc.

Randomization of presentation orders is required for statistical validity but it is also important due to presentation order effects, specifically, first-position order effects. Position order effects occur when the perception of the sample is affected by the position in the presentation sequence that the sample is assessed at. In other words, the first sample is perceived differently than subsequent samples, solely due to its position in the line-up. This so-called first-position effect is quite strong (especially in consumer studies) and the sensory scientist should attempt to mitigate the effect. Randomization with each sample in the first position an equal number of times decreases the effect by spreading it across all samples. A better solution is to serve a dummy sample first followed by the true samples—in this case the panelists are told that they would be served say five samples, but unbeknownst to them the first sample is a dummy and samples two through five are the actual samples. There also seems to be a small but persistent final sample effect.

3.5 Panelist Considerations

3.5.1 Incentives

Some incentive to participate in a sensory study is usually necessary in order to motivate people to volunteer. Sensory specialists should not expect automatic agreement of a person when they are asked to be on a panel and should be realistic about the benefits for that person. "What is in it for me?" is a reasonable question that sensory panel leaders should be ready to answer. In academic settings the days of ordering graduate students to participate are long gone. Likewise, in industry, sensory panel participation should be a volunteer activity. If it is required as a condition of employment (this is not recommended, except in the case noted below), the nature of the participation and the testing must be spelled out at length during the interviews and hiring process, otherwise the voluntary nature of participation is violated.

A guideline for motivating participation is the concept of the token incentive. By "token" we mean that the incentive is just enough to get the person to participate in the evaluation, but not so much that it is the sole reason for the participation. Obviously, if people are paid a great deal they will do just about anything, but an overpaid individual may have little or no motivation to concentrate and work during the evaluation sessions. In other words, they are just "in it for the money." The importance of the token incentive, payment or reward is different in different testing situations. In consumer work, where there is little or no loyalty, commitment or long-term interest in the testing program, the payment is of primary concern. For employees, students, or academic staff who participate in a sensory test, there are other reasons to become involved, such as positive feelings from helping out in the testing program. In some cases and in some cultures, the sense of social responsibility or support for the group effort may be strong enough so that the tokens may be quite minimal.

Common token incentives include snacks or "treats." This can serve as a social or coffee-break time for employees or staff and the opportunity for social interaction may itself become a motivating factor.

Small gifts for repeated testing and free company products are also common incentives. For very high levels of participation, larger gifts or social activities such as a luncheon or a holiday party can recognize the contributions of regular participants. One of us and at least one company that we know of uses a raffle system to entice panelists to attend. An entry is made after each test session. The more tests a person attends the better the chance of winning a prize. This system works well as long as the winners are rotated (you cannot win 2 months in a row) and the sensory professionals themselves are not eligible.

One of the most important incentives for participation is management recognition. When management acknowledges participation in sensory panels as an important contribution to the research effort, recruiting panelists become a great deal easier. Support for sensory evaluation must extend through all levels of management. If only top management supports sensory panel participation, the support quickly becomes "lip service." Supervisors may resent the time employee panelists may spend away from their main job. Thus, it is important to get the cooperation and support of the panelist's direct supervisor as well as all those higher up. An enlightened management will recognize that sensory panel participation enhances job skills, provides a broader motivation for project success, and can serve as a welcome break from routine activities that may enhance overall job performance. It is the job of the sensory professionals to communicate these benefits and to secure management support and to make sure that the supporting attitudes are made known to all potential panelists.

In some companies descriptive panel members are actually additional part-time employees. In this case these employees' only job description is to be panel members. If a sensory specialist decides to employ such descriptive panelists, there must be enough work to keep the panel busy on an ongoing basis. During slow times the panelists may work on re-training exercises or they may be laid off (not a good way to keep the panelists motivated).

3.5.2 Use of Human Subjects

Sensory specialists should be very aware of the health and safety of their panelists. These panelists are human subjects and the specialist should know and follow the guidelines that constrain the use of human subjects. The basis for the guidelines associated with the use of human subjects is the Nuremberg Code of Ethics in Medical Research (United States v. Karl Brandt et al., 1949) and the declaration of Helsinki (Morris, 1966). These guidelines principally state the following:

(1) It is essential the subjects give voluntary consent to participation.
(2) The subject should have the legal capacity to give the consent.
(3) The subject should be able to exercise free power of choice about participating in the study.
(4) The study should yield fruitful results for the good of society.
(5) The researchers should protect the rights and welfare of all the subjects.
(6) The researchers must ensure that the risks to the subjects associated with the study do not outweigh either the potential benefits to the participants or the expected value of the knowledge sought to society.
(7) Above all, the researchers must ensure that each person participating in the study had the right of adequate and informed consent without undue duress.

In legal language, most sensory studies pose no risk "above the ordinary risks of daily life." This includes any inherent risks associated with an individual's chosen occupation (the risks of being an astronaut are greater than those of a college professor). In general, in the United States, sensory testing of foods are often exempt from human subjects oversight scrutiny under the Federal Register (CFR 56:117 § A7 28102). The subjects may be at increased risk if the research plan deviates from the application of accepted and established methods. Physical risks may sometimes be present. For example, food ingredients and additives are sometimes tested during the product development cycle before these ingredients have achieved government approval such as the "generally recognized as safe" (GRAS) list in the United States. Employees who are asked to participate in such tests should be told of all possible risks and as always, participation should be voluntary. Additionally, sensory specialists should be sensitive to psychological risks such as embarrassment when mistakes are made. If panel results are published,

shown or otherwise made available, as in some panelist training and monitoring situations, care must be taken to protect the feelings and, if possible, the identity of the outliers in the data.

In academic settings in the United States, all studies involving human subjects must be reviewed and approved by the particular institution's Human Subjects' Institutional Review Board (Belmont Report, 1979; Edgar and Rothman, 1995). In industrial settings this is not required. However, the ethical sensory specialist will adhere to the principles associated with responsible research involving human subjects (Sieber, 1992).

3.5.3 Panelist Recruitment

The sensory specialist must make sure that the people who are recruited know what is expected of them during the study. It is best to view their participation in the study as a contractual relationship. As much information as possible about time commitment and the product categories should be available to the potential panelists before they commit to the project. Panelists must also be told clearly what they may expect to gain from the study, such as daily treats, money. In most settings, the sensory specialist must be sure that panelists have approval from their supervisors to participate. Additionally, in academic settings in the United States, depending on the specific institution's Human Subjects Institutional Review Board, the sensory specialist may also need to make sure that each panelist voluntarily signs an informed consent form prior to participation in the study.

3.5.4 Panelist Selection and Screening

For certain product categories it may be necessary to have the panelists undergo a medical screening prior to participating in the sensory study. Additionally, the sensory specialist may need to screen the sensory acuity of the potential panelists. However, the specialist should allow some leeway in the sensory deficiencies of the potential panelists. Some people may be very discriminating in general, but have one or two problem areas. Also, many average panelists will improve markedly with training. Therefore it is not necessary to have only the most highly discriminating panelists at the outset of training.

To screen for panelists the sensory scientist should create a battery of tests that are appropriate to the products to be evaluated and the general tasks required of the panelists. If the panelists are only going to be doing discrimination testing then the screening tests should only involve discrimination tasks. On the other hand, if the panelists are going to do scaling tasks then the screening tests may involve both discrimination and scaling tasks. The key to screening, however, is not to over-test panelists before performing true product evaluations. Too many screening tests could decrease the panelists' enthusiasm and motivation when it comes time to do "real" products. Judicious decisions related to the amount of screening needed for a specific study are important.

3.5.4.1 Examples of Screening Tests

The sensory scientist can create a series of discrimination tests differing in difficulty. In other words, the sensory scientist creates a series of product formulations that are more and more difficult to tell apart. Jellinek (1985) discusses how to select panelists using an extensive training course. She required the panelist to meet a stringent series of minimum requirements, before the panelist will be allowed to participate in sensory evaluation studies. These are generally applicable to a broad range of food testing. If the sensory program is more limited in scope, a series of tests may focus on the specific attributes to be encountered in the food products to be tested.

Additionally, it is helpful during panel screening to determine whether the panelist can discriminate the key ingredient flavors and the possible taints (off-flavors) in the product. It is possible to ask the panelists to rank order the intensity of the key ingredient flavors in the product or to rank order increasing levels of taints in the product. Panelists could also be asked to use multiple choice tests to describe aroma, flavor, and mouthfeel characteristics of the products. The sensory specialist can use these data to determine the extent of panel training. Such testing may illuminate areas needing additional work or identify panelists requiring special consideration and training.

If possible, the sensory specialist should recruit two to three times as many persons as needed for the panel. Then rank the panelists' performance on several screening test measures and invite the top performers to participate in the actual panel. A sensory specialist must be very tactful when potential panelists are told that their services will not be needed. When the panelists are informed of their performance on the screening tests, the specialist should use general labels, all of which should be positive in connotation. For example, the group could be divided into good smellers, very good smellers, and excellent smellers, rather than describing groups by using adjectives like "bad" or "poor." It is necessary to be very diplomatic and very careful not to insult people. All potential panelists must be made to feel appreciated, even if they were not invited to participate at this time since they may be recruited for a different study at a later time.

Records of all screening tests should be kept to compare future performance of these and new panelists. It is very possible that the performance of some panelists will improve over time and that of others will get worse. The specialist should plan from the beginning what to do about panel attrition because it will occur. The decision must be made whether panelists will be trained and added over time or whether the final panel size will be smaller than originally planned.

The most important fact to remember is that good panelists are not born but they can be created through the hard work of the panelist and the sensory specialist. Most individuals of average sensory activity can be trained to a level of very high, reliable, and accurate sensory evaluation performance.

3.5.5 Training of Panelists

The amount of training required is dependent on the task and the level of sensory acuity desired. For most descriptive tests extensive and in-depth training is necessary (see Chapter 10). For many discrimination tests, only minimal training is involved. In these cases the panelists are oriented to the task and that is the extent of the training (see Chapter 4).

During the training phase, especially for descriptive panels, the sensory specialist must make the panelists realize sensory work is difficult and requires attention and concentration. During extensive training sessions it is helpful if the panel develops an esprit de corps and this can be facilitated during training by having the panelists work as a team. As mentioned earlier, panelists are easier to train and likely to remain more motivated during the entire study if they feel that the sensory work done by them is valued by management. Attrition and turnover on panels are a factor in all settings. The sensory specialist must plan for this from the first day of recruitment. It is sometimes possible for new trainees to work with experienced panelists, such as people who had been trained for another product category.

3.5.6 Panelist Performance Assessment

The performance of trained panelists used over long periods of time may fluctuate, as the panelists become more or less motivated to participate and to concentrate on the task at hand during evaluation sessions. Also if people do not participate for awhile due to transfers, vacations, leaves-of-absence, etc., their performance may deteriorate and require re-training. Many companies have panelist assessment and reporting programs in place. These can be as simple as plotting the scores given by each panelist against the mean scores for the panel or as elaborate as using multiple assessment programs like those described by Sinesio et al. (1990), Naes and Solheim (1991), Mangan (1992), and Schlich (1996). The Panel Check program (available for free from the European Sensory Network (ESN) website) incorporates most of the above assessment programs in a single simple-to-use software package. We will revisit this issue in Chapter 10.

3.6 Tabulation and Analysis

3.6.1 Data Entry Systems

With the decrease in costs associated with personal computers a number of data entry systems have become very readily available. In this section we are not going to compare the systems currently available,

since these would be obsolete within a few years. We are, however, going to list some principles that should be kept in mind when sensory specialists explore the feasibility of different data entry systems in their facility.

1. The limitations of the computer system should not dictate the form of the test. Before purchasing a system the sensory specialists should be sure that all the tests used in their situation can be programmed with the specific software system.
2. Purchasing online computerized systems requires a careful evaluation of cost-savings in terms of technician time and data entry time "by hand" and the pay-back time versus the overall cost of the system as well as the time needed to become comfortable with using the system.
3. In most situations the testing volume is the primary determinant of the need for automation or direct online entry. In situations where small volumes of many different types of tests are performed, a computerized system may also be useful.
4. The sensory specialist should be aware that less expensive alternatives to online data entry exist: the use of digitizing to enter data or the use of optical scanning.
5. The advantages of computerization of the sensory booth include

   (a) the speed of receiving test results
   (b) a ready interface between the data entry system and statistical and graphing programs
   (c) a reduction in the errors involved in data entry ("key punching")

6. Disadvantages include

   (a) consumers may be unfamiliar with computers and ill-at-ease with using the system. Their concentration may shift to the response system rather than the products
   (b) errors in use may go undetected if data are analyzed "automatically," e.g., without inspection
   (c) computer programs may not be flexible enough to handle variation in experimental designs or requirements for specific scale types

3.7 Conclusion

In many ways the good practice techniques associated with sensory testing are based on common sense. Many of the coding and setting up practices seem very cumbersome at first glance, but the goal is to insure that the specialist always, at all times, knows which sample is in which coded container, because inevitably at some point in a study a sample will be spilled. Sensory specialists should continually ask themselves whether a specific serving container, serving procedure, panelists' recruitment method seem logical and sensible.

Additionally, the use of good practice techniques improves the quality of the tests performed and this in turn will instill client confidence which ultimately leads to increased management respect for the results of sensory tests.

References

Amerine, M. A., Pangborn, R. M. and Roessler, E. R. 1965. Principles of Sensory Evaluation of Foods, Academic, New York, Ch. 6.

Belmont Report. 1979. Ethical Principles and Guidelines for the Protection of Human Subjects Research. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. National Institutes of Health. Office for the Protection from Risks Research. Washington, DC.

Bett, K. L. and Johnson, P. B. 1996. Challenges of evaluating sensory attributes in the presence of off-flavors. Journal of Sensory Studies, 11, 1–17.

Brownlee, K. A. 1957. The principles of experimental design. Industrial Quality Control, 13, 1–9.

Calviño, A., González Fraga, S. and Garrido, D. 2004. Effects of sampling conditions on temporal perception of bitterness in Yerba mate (*Ilex paraguariensis*) infusions. Journal of Sensory Studies, 19, 193–210.

Cardello, A. V. and Segars, R. A. 1989. Effects of sample size and prior mastication on texture judgments. Journal of Sensory Studies, 4, 1–18.

Cochran, W. G. and Cox, G. M. 1957. Experimental Designs. Wiley, New York.

de Wijk, R., Engelen, L. Prinz, J. F. and Weenen, H. 2003. The influence of bite size and multiple bites on oral texture sensations. Journal of Sensory Studies, 18, 423–435.

Edgar, H. and Rothman, D. J. 1995. The institutional review board and beyond: Future challenges to the ethics of human experimentation. The Milbank Quarterly, 73, 489–506

Eggert, J. and Zook K. 2008. Physical Requirement Guidelines for Sensory Evaluation Laboratories, Second Edition. ASTM Special Technical Publication 913. American Society for Testing and Materials, West Conshohocken, PA.