Chapter 2
Psychological Considerations in Sensory Analysis

Jeannine Delwiche

To many, the term "psychology" conjures the image of a distraught patient lying on a couch, telling her most intimate thoughts to a bearded man smoking a cigar and scribbling notes somewhere behind her. "What on earth do interpreted dreams, unhappy childhoods, and envy for certain aspects of male anatomy have to do with the sensory evaluation of dairy products?" you may ask. The answer is, "Not much." When we talk about psychological considerations in sensory analysis, we are not calling upon the ghost of Sigmund Freud, but instead referring back to some of his predecessors and contemporaries up north in Germany: Ernst Weber, Gustav Fechner and Wilhelm Wundt. These men were all pioneers in the area of experimental psychology, a branch of psychology that does not rely upon interviews and introspection, but rather upon the experimental method. Experimental psychology, in essence, does not trust the individual to be able to accurately tell the researcher what features are most important in determining a response. Instead, through careful design and controls, experimental psychology forces the individual to demonstrate what aspects are most important and to more or less "Prove it."

The sub-discipline of experimental psychology known as psychophysics is of greatest relevance to sensory analysis. Fechner, while working in Weber's lab, gave rise to psychophysics with the publication of *Elemente der Psychophysik* (1860). Psychophysics is the area of natural science that deals with sensory physiology and which strives to explain the relationship between sensory stimuli and human responses. A major focus of psychophysics is to discover the relationship between a stimulus (C) and the resulting sensation (R). In its simplest form, this expression may be expressed as a mathematical function (f), $R = f(C)$. Inspired by Fechner's treatise, Wilhelm Wundt is credited with establishing the first laboratory for psychological research. The tools upon which the psychophysicist relied, and often still relies, were measured thresholds and direct scaling, tools that are often used today in sensory analysis. A complete discussion of psychophysics was provided by Amerine et al. (1965). Much of the early work in psychophysics was devoted to discovering how well a

---

J. Delwiche
The Ohio State University, Columbus, OH

S. Clark et al. (eds.), *The Sensory Evaluation of Dairy Products*,
DOI 10.1007/978-0-387-77408-4_2,

person could detect a stimulus. This was sought through the determination of threshold values, which is the minimal quantity of a substance or compound that can be detected, or the boundary at which the subject crosses from "not detecting" to "detecting."

When conducting psychophysics, the researcher begins with an experimental stimulus that can be measured objectively, and such stimuli can range from pure tones of known energy to salt solutions of known concentration. The investigator presents the stimulus in a neutral and repeatable fashion to the subject, and then records the subject's assessment of that stimulus. After multiple presentations and assessments of the test stimuli, often by more than a single subject, the respective responses are analyzed statistically to determine the ways in which the subjects perceive the test stimuli. Similarly, the sensory analyst starts with known products, such as yogurts made at different production sites, presents the products in a neutral and repeatable fashion to the panelist, and then records the panelist's assessment of the products. After multiple presentations and assessments of the products, typically from more than one panelist, the responses are analyzed statistically to determine the product characteristics.

In psychophysics, the goal is to understand how individuals perceive the physical world, whereas in sensory evaluation the goal is to understand the perceptual characteristics of the products. Nonetheless, the tools used in both psychophysics and sensory evaluation are the same and are subject to similar constraints when it comes to best practices.

2.1 Tools of the Trade

The basic tools used by sensory analysts and psychophysicists are: (1) difference tests, (2) ratings, and (3) thresholds. One of the simplest tools utilized by both psychophysicists and sensory analysts are threshold measurements, several types of which have been identified to define more precisely the relationships between the magnitude of a given response and the perceived sensations (Amerine et al., 1965; Meilgaard et al., 1999). There are four types of thresholds (detection, recognition, difference, and terminal) that can be measured, but only two (detection and difference) can be measured with sufficient objectivity to be reliable measures. The easiest threshold to conceptualize is the *detection*, or absolute, *threshold*. It is the smallest amount of a particular stimulus that can elicit a sensation; stimuli of the same type with less intensity do not give rise to sensations. When dealing with taste and smell, the physical intensity is measured by concentration. Thus the threshold for a particular taste or smell is the lowest concentration of a compound that a panelist can distinguish from water (or other solvent). At and above this concentration, the panelist will indicate that a compound is present, while below this concentration the panelist will indicate there is no compound present. Hence, detection thresholds are one way of establishing the relative potencies of different compounds, although caution must be used when making this comparison.

Actual differences in perception across individuals constitute a part of the variability in sensory data that sensory analysts learn to accept and psychophysicists learn to measure. In a study that examined the sensory threshold of off-flavors caused by either proteolysis or lipolysis of milk, 63% of the panelists detected an off-flavor at or below 0.35 mEq of free fatty acids (FFA)/kg milk (Santos et al., 2003). At a FFA concentration of 0.25 mEq FFA/kg milk, only 34% of the panelists could detect the off-flavor (also called rancid off-flavor). As illustrated by this example, the differences in individual thresholds may create a dilemma for milk marketing and quality assurance of fluid milk processors. With a wide range of individual consumer sensory thresholds for rancid off-flavor, where should the acceptance FFA-value be established? Threshold values also vary with testing or serving conditions (Amerine et al., 1965). For these reasons, threshold values are difficult to compare and must be interpreted with caution.

The *recognition threshold* is the level of a stimulus at which the specific stimulus can be recognized and identified. Typically, this level is higher than the detection threshold for the same stimulus. For example, if one was determining the threshold for diacetyl, the concentration at which it was detected would be lower than the concentration at which the aroma would be identified as "buttery." As mentioned above, this sensory measure cannot be made with complete objectivity. The reason has to do with the inability to control for response bias, a topic discussed below.

The *difference threshold* is the extent of change in a stimulus necessary to produce a noticeable difference. The amount of change needed is often referred to as the just-noticeable difference or "*jnd*." The difference threshold is quite similar to the detection threshold, but instead of looking for the lowest intensity that can elicit a sensation, one is determining the lowest increase in stimulation from some base intensity that can elicit a change in sensation. For example, given a baseline concentration of propionic acid, the jnd is the amount of propionic acid that must be added to the baseline concentration before it can be distinguished from the sample containing only the baseline concentration.

A complicating issue with the difference threshold is that the amount of stimulus that must be added to the baseline to be noticeably different increases as the intensity level of the baseline is raised. As an example, consider a room illuminated by candle light with only 10 candles. Let us speculate that the difference threshold is a single candle, and that adding one candle's illumination to the room will increase the illumination by a just-noticeable amount. If we then raise the number of the candles in the room to 100, adding a single candle will no longer raise the illumination level by a noticeable amount. In fact, the just-noticeable difference (jnd) will now be 10 candles. This phenomenon is described by *Weber's law*, which states that the difference threshold divided by the baseline intensity remains constant. Difference thresholds change with stimulus intensity in a predictable way, or stated mathematically

$$\text{Weber's law: } \Delta C/C = k;$$

where C is the absolute intensity of the stimulus, k is a constant (usually between 0 and 1), and ΔC is the change in intensity of the stimulus that is necessary for 1 jnd.

Thus, using our candle illumination example above, we can see that $1/10 = 10/100 = 10\%$. Another way of stating this is that the size of a jnd is a constant proportion of the original stimulus value.

Another practical interpretation of Weber's law indicates that the amount of an added flavor that is just detectable depends on the amount of that flavor that is already present. Knowing k allows the determination of how much added flavor compound is needed for a difference to be noted. Fechner (1860) further refined the relationship between stimulus and response:

$$\text{Fechner's law: } R = k \log C$$

where R is the magnitude of the sensation, k is a constant, and C is the magnitude of the stimulus.

Stevens (1957) described that the perceived magnitude of a response grows as a power function of the stimulus intensity:

$$\text{Stevens' power law: } R = kC^n$$

where R is the response, k is a constant, C is the absolute intensity of the stimulus, and n is the exponent of the power function (a measure of the rate of growth of the perceived intensity, as a function of stimulus intensity).

When n is larger than 1, the perceived sensation grows faster than the stimulus, as is the case for electric shock (3.5) or perception of weight (heaviness) (1.45). When n is smaller than 1, as is the case for many odors, the sensation grows more slowly in relation to the stimulus. A more comprehensive list of power functions is available in Meilgaard et al. (1999). However, just as with thresholds, exponents derived from power laws vary depending upon the subjects making the assessments as well as the methods used to determine them, often making direct comparisons of published values difficult.

The fourth type of threshold that can be measured is the *terminal threshold*, which is the magnitude of a stimulus above which there is no increase in the perceived intensity of the appropriate quality for that stimulus. Often, if the stimulus is increased in intensity beyond this level, pain occurs instead. For example, a solution of sodium chloride can become so concentrated that when it is sipped it not only elicits the sensation of saltiness, but also sensations of burning and/or stinging. The terminal threshold would be the highest concentration of sodium chloride above which there is no increased saltiness, only increased burning and stinging. As is the case with the recognition threshold, this measurement is prone to response bias and thus cannot be established with complete objectivity. There are a variety of procedures that can be used to determine thresholds, the details of which are beyond the scope of this chapter. What is important to note is that all modern assessments of thresholds, including those recommended by ASTM International, avoid single stimulus judgments and otherwise control for response bias.

As mentioned above, response bias interferes with the ability to make objective measurements. When a person is asked to make a single stimulus judgment, such as whether or not an aqueous solution contains a compound or if it is simply water, there are two distinct features that influence their decision: sensitivity and response bias. When measuring a threshold, the researcher is interested only in the sensitivity of the panelist. However, the response of the individual is also influenced by that individual's *response bias*, or that individual's willingness to say, "Yes, I detect something other than water." An individual's response bias can be influenced by a variety of circumstances that are independent from the samples and his or her sensitivity, including emotional state, associated consequences of stating there is a stimulus (will the subject receive payment if she is correct? A shock if he is incorrect?), the percent of time a test stimulus (such as a low concentration of sodium chloride) is presented instead of a control stimulus (such as water), distractions within the test environment, etc. As the interests of both psychophysicists and sensory analysts are inclined toward measures of sensitivity, intended to assess sensory systems or product differences, modern sensory procedures are designed to eliminate response bias. To this end, a forced-choice difference test (discussed below) is typically incorporated into the determination of thresholds. In other words, rather than relying upon a panelist to state that he/she can detect a compound in solution, the panelist is asked to *demonstrate* his/her ability to detect it by selecting the sample that contains the compound from a set that contains both blanks and the compound in solution. In each sample set the concentration of the compound is increased until the panelist can reliably select the sample with the concentration over samples that do not contain any compound (the blanks). In other words, instead of relying on the panelist to introspect upon whether or not a compound is present, the subject is asked to *prove* he/she can detect it.

As mentioned above, response bias cannot be eliminated from the measurement of recognition and terminal thresholds, which makes them far less reliable measures than detection and difference thresholds. When measuring a detection threshold, the panelist is challenged to select which unknown in a set of blanks and test stimuli contains the compound. When measuring a difference threshold, the panelist is asked to select which unknown in a set of baseline concentrations and test concentration contains *more* of the compound. Both of these tasks are *forced-choice difference tests*. Regardless of whether or not the panelist would be inclined to call all the samples the same or all the samples different from one another, he/she is forced to select a single sample, eliminating the individual's response bias from the task. It is not possible to set up such a force-choice situation for the measurement of either a recognition threshold or a terminal threshold. Recognition relies upon the individual's willingness to say that he/she recognizes the stimulus, which is his/her response bias. It is unfair to present a set of blanks and test stimuli and then ask the panelist to indicate which he/she recognizes – in actuality he/she may recognize none of them. Furthermore, it is unfair to ask a panelist to ignore pain when tasting extremely high concentrations of compounds in the course of measuring a terminal threshold.

2.2 Neutrality Is Key

Regardless of the sensory tool used (difference test, ratings, or thresholds), neutrality of sample presentation is key. This is because when measuring subtle differences between test stimuli, the panelist will draw upon all available cues in making his/her assessments. Sensory evaluation tradition suggests that samples be labeled with neutral, randomly generated 3-digit numbers. Numbers with inherent meaning should be avoided (i.e., 666, 911, local area code, etc.). While it is not entirely necessary to use such labels, they are among the safest choices. Labels should not imply order or sequences, nor should they suggest quality, thus labels such as A, B, C or 1, 2, 3 are particularly problematic. Two-digit numbers are often associated with sports figures and are generally less desirable for labeling samples. All labels should be generated in the same fashion, either on sticker labels or written directly on cups. All labels should be printed with the same font and style, or all written in the same hand-writing, and all should be of the same color.

Other aspects of sample presentation should also be neutral. All samples need to be served at the same volume and same temperature. All samples should be served in identical neutral containers. Crushed cups and dented lids should not be used. When presented to the panelists, all samples should be presented with labels facing forward. Careful presentation is necessary to ensure that assessments are based only upon the characteristics of the samples themselves rather than upon extraneous cues.

2.3 Perception Is More than a Sum of Its Parts

When asking panelists to assess dairy products, it is important to remember that the perceptual experience that occurs when a sample is placed in the mouth is a gestalt – a unified whole that cannot be derived from the summation of its component sensations. Not only are sensations of taste and smell elicited, but a variety of other sensory systems are also activated including sight, temperature, and texture. These sensations interact with one another and create the gestalt experience of flavor. Furthermore, it is simply not possible for a panelist to ignore a particular sensation while assessing others, even if the panelist attempts to comply with such instructions. A trained panelist may learn to separate the different aspects of the unified experience, but these sensations interact in the creation of the whole and the alteration of the components occurs before the panelist has the chance to disentangle them.

For example, taste and smell interact. Increasing the concentration of odor compounds typically increases ratings of taste intensity and increasing the concentration of taste compounds generally increases ratings of smell intensity (Bonnans and Noble, 1993; Frank et al., 1989; Murphy and Cain, 1980; Murphy et al., 1977; Philipsen et al., 1995). Differential effects due to taste and odor qualities further complicate the issue. For example, the addition of sucrose to fruit juices not only increases sweetness and fruit odor, but also decreases bitterness, sourness, and unripe odor (von Sydow et al., 1974). Color also interacts with taste and smell, wherein ratings of taste, smell, and flavor generally increase as color intensifies (DuBose et al., 1980; Johnson et al., 1982; Johnson and Clydesdale, 1982; Johnson et al., 1983 Norton and Johnson, 1987; Teerling, 1992). Additionally, appropriately colored foods and beverages are identified correctly more often than uncolored and/or inappropriately colored items (DuBose et al., 1980; Hall, 1958; Moir, 1936; Philipsen et al., 1995; Stillman, 1993; Teerling, 1992). This is likely due to individuals associating certain flavors with specific colors and when the colors are altered, identification becomes more difficult.

Texture impacts the perception of dairy products both directly and indirectly. How thick or thin, smooth or lumpy, crumbly or springy, etc., all impact assessments of the product. However, texture characteristics also control the concentration of taste and smell compounds released as well as the rate at which they are released (Overbosch et al., 1991). Increasing a product's thickness slows the diffusion of components to the sensory receptors while decreasing the thickness will increase the rate of diffusion. This means that two items with identical amounts of taste and smell compounds but different body/texture will differ in perceived taste and smell intensities. Additionally, the thicker-textured item will take longer to reach its peak taste and odor intensities and peak intensities will typically be lower than those of the thinner product.

Similarly, temperature impacts the perceived taste and smell (Delwiche, 2004). As a product is warmed, there is an increase in volatile components being released from it and correspondingly odor intensity becomes stronger. Temperature itself can elicit taste sensations (Cruz and Green, 2000); thus, changing a product's temperature will alter its taste intensities. In addition, increasing product temperature can also decrease product thickness, resulting in the concomitant increases of taste and smell intensities as described above.

Appearance, aroma, flavor, body and texture interactions are real and complex, and beyond conscious awareness and control. For these reasons, it is simply not possible for panelists to ignore specific perceptual features when making their assessments. If the researcher is interested in flavor differences but the products differ in texture, the researcher must realize that the flavor assessments will be impacted by the texture differences unless those differences are somehow eliminated. Nor is it possible for a panelist to ignore a temperature difference between samples if they are served at different temperatures. In fact, if a difference test is being conducted, sample temperature may be one of the ways that panelists differentiate the products.

2.4 Sensory Analysis: A True Science

From the discussion just presented, it should be clear that sensory analysis, like psychophysics, is a natural science. Like all natural sciences, measurements of sensory characteristics of foods or beverages can and should be taken carefully. When done properly, sensory information can provide great insight into the world. When measures are undertaken poorly they do more to mislead than to inform. Careful controls must be implemented and followed when conducting sensory analysis, including (1) neutrality in the presentation of samples, (2) elimination of response bias, and (3) use of methods that require panelists to demonstrate their ability rather than relying upon self-reports. Failure to adhere to any of these controls diminishes the value of the resulting sensory data. By contrast, determining appropriate controls and ensuring they are in place will result in reliable and useful information about foods and beverages which no instrument can measure – their perceptual characteristics.

References

Amerine, MA, Pangborn, RM and Roessler, EB. 1965. Principles of Sensory Evaluation of Food. Academic Press, New York. 602 pp.

Bonnans, S and Noble, AC. 1993. Effects of sweetener type and of sweetener and acid levels on temporal perception of sweetness, sourness and fruitiness. Chemical Senses 18: 273–283.

Cruz, A, & Green, BG. 2000. Thermal stimulation of taste. Nature 403 (6772), 889–892.

Delwiche, JF. 2004. The impact of perceptual interactions on perceived flavor. Food Quality and Preference 15: 137–146.

DuBose, CN, Cardello, AV and Maller, O. 1980. Effects of colorants and flavorants on identification, perceived flavor intensity, and hedonic quality of fruit-flavored beverages and cake. Journal of Food Science 45: 1393–1399, 1415.

Fechner, GT. 1860. Elemente der Psychophysik. Leipzig, Breitkopf und Härtel.

Frank, RA, Ducheny, K and Mize, SJS. 1989. Strawberry odor, but not red color, enhances the sweetness of sucrose solutions. Chemical Senses 14: 371–377.

Hall, RL. 1958. Flavor study approaches at McCormick and Co., Inc. In: AD Little, editor. Flavor Research and Food Acceptance. Reinhold, New York. 224–240.

Johnson, JL and Clydesdale, FM. 1982. Perceived sweetness and redness in colored sucrose solutions. Journal of Food Science 47: 747–752.

Johnson, JL, Dzendolet, E and Clydesdale, FM. 1983. Psychophysical relationships between sweetness and redness in strawberry-drinks. Journal of Food Protection 46: 21–25.

Johnson, DL, Dzendolet, E, Damon, R, Sawyer, M and Clydesdale, FM. 1982. Psychophysical relationships between perceived sweetness and colour in cherry-flavoured beverages. Journal of Food Protection 45: 601–606.

Meilgaard, M, Civille, G and Carr, B. 1999. Sensory Evaluation Techniques (3rd Edition). CRC Press, New York.

Moir, HC. 1936. Some observations on the appreciation of flavor in foodstuffs. Chemistry and Industry 14: 145–148.

Murphy, C and Cain, WS. 1980. Taste and olfaction: Independence vs. interaction. Physiology & Behavior 24: 601–605.

Murphy, C, Cain, WS and Bartoshuk, LM. 1977. Mutual action of taste and olfaction. Sensory Processes 1: 204–211.

Norton, WE and Johnson, FP. 1987. The influence of intensity of color on perceived flavor characteristics. Medical Science Research: Psychology & Psychiatry 15: 329–330.

Overbosch, P, Afterof, WGM and Haring, PGM. 1991. Flavor release in the mouth. Food Reviews International 7: 137–184.

Philipsen, DH, Clydesdale, FM, Griffin, RW and Stern, P. 1995. Consumer age affects response to sensory characteristics of cherry flavored beverage. Journal of Food Science 60: 364–368.

Santos, MV, Ma, Y, Caplan, Z and Barbano, DM. 2003. Sensory threshold of off-flavors caused by proteolysis and lipolysis in milk. Journal of Dairy Science 86: 1601–1607.

Stevens, SS. 1957. On the psychophysical law. Psychological Review 64: 153–181.

Stillman, JA. 1993. Color influences flavor identification in fruit-flavored beverages. Journal of Food Science 58: 810–812.

Teerling, A. 1992. The colour of taste. Chemical Senses 17: 886.

von Sydow, E, Moskowitz, H, Jacobs, H and Meiselman, H. 1974. Odor-taste interaction in fruit juices. Lebensmittel-Wissenchaft und -Technologie 7: 9–16.

http://www.springer.com/978-0-387-77406-0

The Sensory Evaluation of Dairy Products
Clark, S.; Costello, M.; Drake, M.; Bodyfelt, F. (Eds.)
2009, XV, 576 p., Hardcover
ISBN: 978-0-387-77406-0