CHAPTER 3

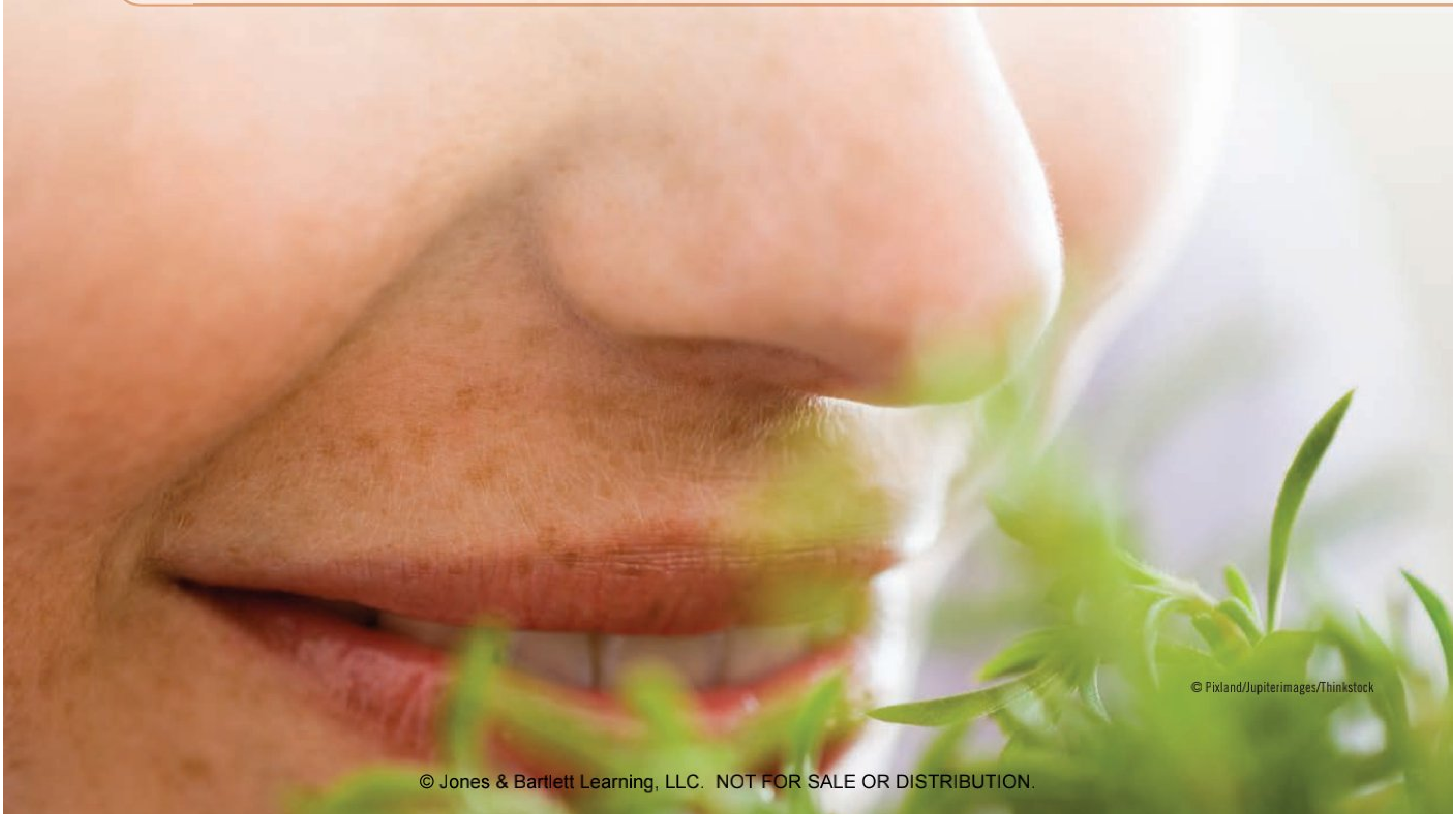
Sensory Evaluation

Sung Eun Choi, PhD, RD

Chapter Objectives

THE STUDENT WILL BE EMPOWERED TO:

- Identify the sensory characteristics of food.
- Discuss the factors affecting the outcomes of sensory evaluation.
- Demonstrate an understanding of the process for sensory evaluation tests.
- Formulate an effective sensory evaluation strategy by selecting appropriate test design, panelists, and instruments.
- Discuss how to analyze and interpret the sensory data and recognize specific methodological advances related to sensory evaluation.



© Pixland/Jupiterimages/Thinkstock

Introduction

Determining how food products affect consumers' senses is one of the most important goals of the food industry. It also is a primary concern for nutritionists and dietitians who develop healthier recipes. Because our five senses act as the gatekeeper of our bodies, the benefits of healthy food will be reaped only if our senses accept it. Therefore, consumer reaction, as perceived by the five senses, is considered a vital measure of food development. Because no apparatus can substitute for the senses in evaluating food, humans are used as test subjects. Such studies are becoming more prevalent, despite the potential biases of humans and the costs involved.

sensory evaluation The scientific measurement method of food quality based on sensory characteristics as perceived by the five senses.

Sensory evaluation is a scientific method that evokes, measures, analyzes, and interprets responses to products, as perceived through the senses of sight, smell, touch, taste, and sound.[1] This widely accepted definition is used by sensory evaluation committees within various professional organizations, including the Institute of Food Technologists and the American Society for Testing and Materials.[2] Like other scientific methods of taking measurements, sensory evaluation is concerned with precision, accuracy, and sensitivity and with avoiding false-positive results.[3] Reliable sensory evaluation is based on the skill of the sensory analyst in optimizing four factors: definition of the problem, test design, instrumentation, and interpretation of the results:[4,5]

- *Definition of the problem:* The item to be measured must be defined precisely.
- *Test design:* Not only must the design take into account unknown sources of bias, but it also must minimize the amount of testing required to produce the desired accuracy of the results.
- *Instrumentation:* The panelists must be selected and trained to give a reproducible result.
- *Interpretation of the results:* The analyst should select appropriate statistics based on the correct statistical assumptions and draw only those conclusions that are supported by the data.

The benefits of well-performed sensory evaluation can be realized in many ways; however, if the sensory analyst fails to optimize one of the four factors, much time and money is wasted. For effective sensory evaluation, the analyst should duly recognize the purpose of the study, select the appropriate experimental design, use panelists who fit the purpose, choose the proper method for preparing and presenting the samples, and analyze the data correctly. A sensory researcher should always consider whether the method is appropriately implemented and whether errors have been introduced at any stage of the experiment.

In this chapter, the principles of sensory evaluation will be introduced. Examples will be provided to demonstrate the use of sensory evaluation techniques and the application of the results toward developing and modifying food recipes.

The Human Senses

The characteristics of food are perceived by the five senses: sight, smell, taste, sound, and touch.

Sight

The eyes perceive the initial quality of food, receiving such information as color, size, shape, texture, consistency, and opacity. Light entering the lens of the eye is focused on the retina, where the rods and cones convert it to neural impulses that travel to the brain via the optic nerve.[5] Perception by the visual system of light of wavelengths 400–500 nanometers (nm; blue), 500–600 nm (green and yellow), and 600–800 nm (red) is commonly expressed in terms of the hue, value, and chroma of the Munsell color system.[5]

Color may accurately indicate ripeness, strength of dilution, and the degree to which the food has been heated. Color is used to evaluate a food's desirability and acceptability. Greenish bananas, burnt meat, and dark brown avocado send visual signals that can change a person's choices. Color often triggers certain expectations in the mind; for example, the creamy color of vanilla ice cream evokes an expectation of richness.

But, color can be deceiving. The quality of food can be masked by changes in color. For instance, if yellow coloring is added to a food without actual fat having been added, the quality of low-fat products can be improved. Color changes alone can increase a food's acceptability considerably.

Even small visual details such as adjacent or background colors and the relative sizes of areas of contrasting color can affect a consumer's perception. Parameters of size and shape, such as width, length, thickness, particle size, geometric shape, and distribution of pieces, also provide information on food quality. The dullness, shininess, smoothness, or roughness of a surface and the clarity of liquids evoke preconceptions about the food. See **Special Topic 3.1** for more information on some of the latest trends chefs use to create dishes that delight the senses.

Special Topic 3.1

Molecular Gastronomy

Sarah Churchill

The concept of molecular gastronomy was founded by Hervé This, editor at *Pour la Science*, and Nicholas Kurti, a low-temperature physicist. They coined the term **molecular gastronomy** to encompass all the physical and chemical changes that occur during food production and cooking. In an effort to identify the methods for creating the best flavor and texture, they compared existing recipes with old proverbs and old wives' tales.[1] In 1992, they established the first international Workshop on Molecular and Physical Gastronomy.[2] The workshop still takes place today; however, the field of molecular gastronomy has moved from the scientific realm into the media limelight, becoming a cultural phenomenon.

molecular gastronomy Form of modern gourmet cooking that seeks to identify all the physical and chemical changes that occur during food production and cooking to create the best flavor and texture.

Today, popular culture uses the term *molecular gastronomy* to refer to the way modern chefs innovate with ingredients and techniques to bring excitement to the dining experience. The practice of molecular gastronomy also has been referred to as *culinary constructivism*, *experimental cuisine*, *molecular cooking*, *modernist cuisine*, *culinary deconstructivism*, and *progressive cuisine*. For example, the modern chefs Ferran Adriá of el Bulli, Heston Blumenthal of The Fat Duck, and Grant Achatz of Alinea are attempting to understand the chemical and physical nature of cooking as well as the best ingredients and techniques to improve upon traditional methods. They accomplish these goals through the deconstruction of certain recipes, the transformation of the physical states of foods (i.e., gases, liquids, solids), and the use of different equipment and ingredients to change cooking methods.[3]

Such chefs do not necessarily ascribe their innovations to molecular gastronomy, however. Adriá and Blumenthal say they go beyond the mere scientific exploration promoted by This and Kurtis. They also dispute the claim that the cuisine they are creating is pure novelty. They explain that they view food with a certain openness and embrace new ingredients and cooking methods without forsaking tradition. They affirm that their methods help them realize the full potential of each ingredient by questioning traditions about the best way to cook a food and finding the best way to maintain flavor. The result is a unique cuisine that gives consumers a chance to engage and confuse all of their senses.[4]

What innovations are used to create this bold new cuisine? Modern chefs use centrifuges, syringes, freeze dryers, blast chillers, and carbon dioxide dispensers. Canapés can be created from centrifuged frozen peas to create a buttery spread. A syringe can even be used to inject a small amount of unexpected flavor. Sometimes molecular gastronomy turns a traditional dish on its head. For example, instead of pasta with grated cheese, how would you like cheese with grated pasta? Parmesan noodles can now be made from boiling the cheese, pressing out the water, and passing it through a pastry bag; the cheese noodles can be topped with grated freeze-dried pasta.[5]

© Thomas_EyeDesign/iStockphoto.com

The kitchen of a chef cooking using molecular gastronomy techniques may look more similar to a high-end chemistry or physics laboratory than the kitchen of the local greasy spoon. A fusion of chemistry and cooking, molecular gastronomy seeks to transform common cooking ingredients into novel forms through the use of new technologies and techniques.

New ingredients at the modern chef's disposal include xantham gum, alginate, calcium salts, soy lecithin, agar-agar, and liquid nitrogen. These ingredients are becoming as prevalent in the modern chef's pantry as spices. They can be used to make gels, foams, and spheres without adding unwanted flavor.[6] Liquid nitrogen can be used to freeze something immediately and even make it shatter. Soy lecithin is crucial in making foams because it helps emulsify and hold ingredients together. Spheres resemble the texture and appearance of caviar but can be made using any liquid.

With the help of new technology and new ingredients, these chefs are able to create unique, artistic dishes. Chefs want to amaze consumers, and they really have. Whatever this cuisine should be called, it has become very popular. Of the 50 best restaurants in the world, the top 3 are associated with molecular gastronomy. Diners are becoming more experimental in their choices, even bravely trying tobacco- or crab-flavored ice cream. Innovation in cuisine is always progressing. With increased knowledge and technology, we'll just have to wait to see what comes next!

References

1 Molecular gastronomy: Food science raises the culinary bar. *Env Nutri.* 2010;33(9):7.
2 Blanck JF. Molecular gastronomy: Overview of a controversial food science discipline. *J Agr Food Inf.* 2007;8(3):77–85.
3 Lanchester J. Incredible edibles: The mad genius of "modernist cuisine." *The New Yorker.* 2011 March:87, 64.
4 Adria F, Blumenthal H, Keller T, McGee H. Statement on the 'new cookery.' *The Observer.* 2006. Available at: http://www.guardian.co.uk/uk/2006/dec/10/foodanddrink.obsfoodmonthly. Accessed June 28, 2012.
5 Adler J. Extreme cuisine. *Smithsonian.* 2011;42:60–66.
6 Ehrenberg R. What's cooking? *Science News.* 2008;173(13):202–203.

Smell

olfactory Relating to the sense of smell.

Our sense of smell, or **olfactory** sense, also contributes to our evaluation of food quality. The volatility of odors is related to temperature. Because only volatile molecules, in the form of gas, carry odor, it is easier to smell hot foods than cold ones. For example, hot tea is much easier to detect than iced tea, and the odor of a baked item is more intense than that of ice cream.

Lighter molecules that can become volatile are detected by the olfactory epithelium in the nasal cavity through one of two pathways: (1) directly through the nose or (2) after entering the mouth and flowing retro-nasally, or toward the back of the throat and up into the nasal cavity.[6] If you drink a carbonated beverage and laugh unexpectedly, you may experience the tingling of bubbles in the nose, showing how the mouth and nose are connected and how molecules can reach the olfactory epithelium by either route.

The gradual decrease in the ability to distinguish between odors over time is called **adaptation**. Adaptation occurs to prevent sensory overload. Dairy farmers who are exposed daily to the smell of manure will gradually become unaware of it, whereas visitors to the farm may be taken aback by the smell. Human subjects have varying sensitivities to odors, depending on hunger, satiety, mood, concentration, presence or absence of respiratory infections, and gender (e.g., women who are menstruating or are pregnant may perceive odors differently).[7] Because different people perceive a given odorant differently, identifying a new odor from a food product requires as large a panel as possible to get valid results.

adaptation The gradual decrease in the ability to distinguish between odors over time.

gustatory Relating to the sense of taste.

taste buds The small parts of gustatory and supportive cells; usually found on the upper surface of the tongue.

papillae Rough bulges or protuberances in the surface of the tongue, some of which contain taste buds.

Taste

Taste, or the perception of **gustatory** input, is the most influential factor in a person's selection of a particular food. For a substance to be tasted, it should be dissolved in water, oil, or saliva. Taste is perceived by the **taste buds** (see **Figure 3.1**), which are primarily on the surface of the tongue, by the mucosa of the palate, and in areas of the throat.

In the middle of each taste bud lies a pore, where saliva collects. When food enters the mouth, bits of it are dissolved in the saliva pools and come into contact with cilia, small hairlike projections, from the gustatory cells.[6] The gustatory cells signal to the brain through cranial nerves. The brain, in turn, translates the nervous electrical impulses into sensations that people recognize as "taste."[6]

Taste buds are found in the **papillae** of the tongue. Two types of papillae contain taste buds. The mushroom-like fungiform papillae on the sides and tip of the tongue generally contain taste buds, and the circumvallate papillae (elevated, large papillae in the form of a "V" toward the back of the tongue) always contain taste buds.[8] As people get older, the original 9,000 to 10,000 taste buds begin to decrease in number, so that people older than age 45 often seek more spices, salt, and sugar in their food.[6]

© F.C.G./ShutterStock, Inc.

FIGURE 3.1 On the tongue, the majority of the taste buds sit on raised protrusions of the tongue surface called papillae. On average, the human tongue has 2,000–8,000 taste buds.

Genetic Variation

Individual variation in taste likely has a genetic component. Studies have demonstrated a link between the ability to taste bitter thiourea compounds and a newly discovered taste receptor gene, *TAS2R38*.[9] Thus, the ability to taste these bitter compounds—phenylthiocarbamide (PTC) or the safer, chemically related compound 6-n-propylthiouracil (PROP)—may be used as a phenotypic marker for genetic differences in perceptions of taste.[10–12] In the United States, the frequency of nontasters is estimated to be 20% to 25% of the population.[13] The frequency varies by gender and race.[13,14]

One factor that may explain variation in taste and the perception of physical sensations is the anatomy of the anterior portion of the tongue. For example, PROP tasters have the most fungiform papillae (FP).[13,15–17] **Special Topic 3.2** provides more information on how genetic variations may influence how we perceive food.

Beyond genetics, variation in taste perception also depends on how perceptible sweet, fatty, and bitter components are in foods and beverages. It also depends on the value a consumer places on other factors, such as health and convenience.[18,19]

Special Topic 3.2

Nutrigenomics

Jill M. Merrigan

nutrigenomics Study of how food affects our genes and the way our bodies respond to nutrients.

Nutrigenomics studies the naturally occurring compounds in the foods and how they affect our bodies based our individual genetic differences.[1] "There is good evidence that nutrition has significant influences on the expression of genes, and, likewise, genetic variation can have a significant effect on food intake, metabolic response to food, individual nutrient requirements, food safety, and the efficacy of disease-protective dietary factors," explains nutrigenomic researcher L. R. Ferguson.[2]

The existence of a particular gene or mutation in many cases indicates a predisposition to a particular disease. Once genetic predisposition has been established, determining whether the disease will progress can be investigated by examining the relationship between the human genome and environmental and behavioral factors. The study of nutrigenomics looks at the expression of the genome with regard to nutrition. According to researcher M. Nathaniel Mead, "although genes are critical for determining function, nutrition modifies the extent to which different genes are expressed, and thereby modulates whether individuals attain the potential established by their genetic background."[3] Today, scientists are exploring nutrigenomics to determine how nutrients may be able protect the genome from damage.

Through the study of nutrigenomics, it may become possible to develop dietary interventions based on an understanding of an individual's nutritional requirements, nutritional status, and genotype. Studying how foods affects individual genes and genotypes will make it possible to design "personalized nutrition" plants that may prevent and cure chronic disease.

Studies have been completed on humans, animals, and cell cultures that reveal that macronutrients (fatty acids and proteins), micronutrients (vitamins), and naturally occurring phytochemicals (such as flavonoids, carotenoids, coumarins, and phytosterols) regulate gene expression in various ways. Micronutrients and bioreactive chemicals in foods are involved in metabolic reactions that determine hormonal balances and immune competence, as well as detoxification processes. Additionally some biochemicals found in foods, such as genistein and resveratrol, act as transcription factors, and therefore alter gene expression. Signal transduction pathways and chromatin structures are altered by other biochemicals, such as choline, and therefore indirectly affect gene expression.[3]

One example of nutrigenomics is folate and the gene for MTHFR (methylenetetrahydrofolate reductase). MTHFR has a role in supplying methionine. Methionine plays a central role in certain metabolic pathways, including those involved in the production of neurotransmitters and in the regulation of gene expression. Folate is required for MTHFR to function efficiently. MTHFR has a common polymorphism that leads to two forms of protein: the reference version (C), which functions normally, and the thermal-labile version (T), which has reduced activity. When individuals have two copies of the reference sequence gene (CC), they have normal folate metabolism. However, individuals who have two copies of the reduced version (TT) and low dietary folate accumulate homocysteine and have less methionine. This combination puts them at an increased risk for vascular disease and premature cognitive decline. By making these connections, individuals with the unstable (TT) genes can take folic acid supplements or increase their folate from food sources to metabolize excess homocysteine and restore their methionine levels to normal.[1]

Additional studies have determined that there are nine key nutrients that may have an influence on genomic integrity in a handful of ways. Six of these nutrients—folate, vitamin B_{12}, niacin, vitamin E, retinol, and calcium—are associated with a reduction in DNA damage. The other three—riboflavin, panthenic acid, and biotin—are associated with an increase in DNA damage similar to that seen upon occupational exposure to genotoxic and carcinogenic chemicals. This suggests that nutritional deficiency or excess can lead to DNA damage as damaging as that seen with exposure to environmental toxins. Other nutrigenomic studies have shown that many antioxidant nutrients and phytochemicals enhance DNA repair and reduce oxidative DNA damage.[3]

Recent research indicates that nutrigenomics may have the potential to prevent, mitigate, and treat chronic diseases and certain cancers by making small but highly useful changes to an individual's diet. In the future, scientists, doctors, and dieticians may be able to move forward to be able to identify a patient's DNA profile for a specific disease and ultimately be able to shape a diet that will reduce their chances of developing that disease. Nutrigenomics may be the answer to the obesity epidemic and improve the way individuals age with better bone and brain health; it may also decrease the risk of developing certain cancers.

References

1 Astley SB. An introduction to nutrigenomics developments and trends. *Genes Nutri.* 2007;2(1):11–13.
2 Ferguson LR. Nutrigenomics: Integrating genomic approaches into nutrition research. *Mol Diagn Ther.* 2006;10(2):101–108.
3 Mead MN. Nutrigenomics: The genome–food interface. *Env Health Persp.* 2007;115(12):A582–A589.

Basic Components of Taste

For many years, four basic tastes were recognized: sweet, salty, sour, and bitter. A fifth, umami, was added more recently. These tastes can be characterized as follows:

- *Sweet:* Substances that produce sweet taste include sugars, glycols, alcohols, aldehydes, and alternative sweeteners.[20]
- *Salty:* The salty taste comes from ionized salts, such as the ions in sodium chloride (NaCl) or other salts found naturally in some foods.
- *Sour:* The sour taste comes from the acids found in food. It is related to the concentration of hydrogen ions (H^+) that are found in the natural acids of fruits, vinegar, and certain vegetables.
- *Bitter:* Bitterness is imparted by compounds such as caffeine (tea, coffee), theobromine (chocolate), and phenolic compounds (grapefruit).[6] Many bitter substances are alkaloids that often are found in poisonous plants.
- *Umami:* This is a most recently defined component of taste, which was identified from a study of seaweed broth.[8] **Umami** is a Japanese word meaning "delicious"—it is evoked by glutamate compounds, which are commonly found in meats, mushrooms, soy sauce, fish sauce, and cheese. Some taste experts do not recognize umami as a taste at this time.

Flavor

Whereas taste relies on the sensation produced through the stimulation of the taste buds, flavor is a broader concept. **Flavor** is the combined senses of taste, aroma, and mouthfeel. **Mouthfeel** encompasses textural and chemical sensations such as astringency, spice heat, cooling, and metallic flavor.

Among the flavor components, aroma is especially important; it provides approximately 75% of the impression of flavor.[21] To get an idea of how the ability to smell affects the perception of flavor, pinch your nose and begin to eat a certain flavor of jellybean. Then, as you are chewing, unpinch your nose: You will clearly sense the difference between when the nose is pinched and unpinched. Suppose you are eating a buttered popcorn–flavored jellybean. While pinching your nose, you can only perceive sweetness, but as soon as you unpinch your nose you can recognize the buttered popcorn flavor. As another example, consider when you have a cold with a badly stuffed-up nose. Everything tastes different. (This is why pinching people's nostrils shut is helpful in mitigating the flavor of an unpleasant medicine.)

Gastronomy Point

Monosodium Glutamate Monosodium glutamate, better known by its abbreviation MSG, is one substance that contributes to the perception of umami.

umami Taste category based on glutamate compounds, which are commonly found in meats, mushrooms, soy sauce, fish sauce, and cheese.

flavor The combined sense of taste, odor, and mouthfeel.

mouthfeel The way that a particular type of food feels in the mouth.

Innovation Point

A New Taste? Research is underway for a sixth taste component relating to our perception of fats. Genetic variation in "fat taste" may help explain our food choices and dietary habits, which, in turn, could influence our nutritional and health status.

Sound

Sound is another sense used in evaluating food quality. Sounds such as sizzling, crunching, popping, bubbling, squeaking, dripping, exploding, and crackling can communicate much about a food. Most of these sounds are affected by water content; thus, their characteristics indicate a food's freshness and ripeness.[6]

Sound is detected as vibrations in the local medium, usually air. The vibrations are transmitted via the small bones in the middle ear to create hydraulic motion in the fluid of the inner ear, the cochlea. The cochlea is a spiral canal covered in cilia that, when agitated, sends neural impulses to the brain.[5]

Touch

texture The sensory manifestation of the structure or inner makeup of products in terms of their feel as measured by tactile nerves on the surface of the skin of the hands, lips, or tongue.

descriptive panel A panel commonly used to determine differences between food samples. The descriptive panelist is experienced in the type of food being tested and receives extensive training prior to the testing.

consumer panel A panel selected from the public according to the demographics necessary to taste test a product.

The sense of touch delivers impressions of a food's texture to us through oral sensations or the skin. Texture is a very complex perception: The first input is visual; second comes touch, either directly through the fingers or indirectly via eating utensils; the third is the feeling in the mouth (mouthfeel), as detected by the teeth and tactile nerve cells on the tongue and palate. **Texture** is the sensory manifestation of the structure or inner makeup of products in terms of their reactions to stress, which are measured as mechanical properties (such as hardness/firmness, adhesiveness, cohesiveness, gumminess, springiness/resilience, and viscosity) by the kinesthetic sense in the muscles of the hands, fingers, tongue, jaw, or lips.[5] Texture also includes tactile feel properties, which are measured as geometric properties (i.e., grainy, gritty, crystalline, flaky) or moisture properties (i.e., wetness, oiliness, moistness, dryness) by the tactile nerves in the surface of the skin of the hands, lips, or tongue.[5] The greater surface sensitivity of the lips, tongue, face, and hands makes easy detection of small differences in particle size and thermal and chemical properties possible among food products.

Variables Controlled During Sensory Evaluation

During sensory evaluation, panelists are typically seated at tables, cubicles, or booths, and the food is presented in a uniform fashion. To obtain valid, reproducible results during a sensory evaluation, the environment in which the sensory panel evaluates foods or beverages should be carefully controlled, as should variables pertaining to the panelists. This section discusses the many variables that should be considered when designing a sensory evaluation test.

Panel Management

Two general types of panels are used in sensory evaluation. A **descriptive panel** is commonly used to determine differences between food samples. The descriptive panelist is experienced in the type of food being tested and receives extensive training prior to the testing. A **consumer panel** is selected from the public according to the demographics necessary to taste test a product.

Panel Selection

When assembling a panel, it is preferred to use an equal number of men and women. The age distribution of the panel should also be considered because it may affect test results.[6] The sensory analyst must recruit the people who can make a reliable commitment of time and who also know what is expected of them during the test. General taste panels usually consist of people who meet the following criteria:

- They are in good health and free of illness related to sensory properties, such as chronic colds, food allergies, or diabetes.
- They are nonsmokers (smoking can dull olfactory and gustatory sensations).

- They are not color blind.
- They have no strong likes or dislikes for the food to be tested.

Panelist Preparation

The level of training for descriptive panels and consumer panels is quite different, given the differences in purpose of the evaluations in which they participate.

Descriptive Panels

Because the investment in a descriptive panel is large in terms of time and human resources, it is wise to conduct an exhaustive screening process rather than train unqualified panelists.[2] If the ability to detect subtle differences is essential, the sensory analyst may need to screen the sensory acuity of potential panelists on key properties of the product(s) that will be tested.

Descriptive panels can be selected through a series of tests that may include a set of prescreening questionnaires, a set of acuity tests, a set of ranking/rating tests, and a personal interview. However, it is not necessary to have only the most highly discriminating panelists because the average panelists will improve markedly with training and some people may be very discriminating in general but just have one or two problem areas.[2]

The amount of training required is determined by the task and the level of sensory sensitivity desired. For most descriptive panels, expensive and in-depth training is necessary.[2] During the training, the trainer must make sure the panelists realize that sensory testing work is difficult and requires attention and concentration. If team spirit can be developed by the panelists during the training sessions, this will smooth the way for the main evaluation and facilitate panelist performance. The performance of trained panelists used over long periods of time may fluctuate because of a loss of focus and a lack of motivation during the evaluation sessions.

Consumer Panels

In contrast to descriptive panels, consumer panels typically require a larger number of panelists and may range from 200 to 500 people (see **Figure 3.2**). Consumer panelists can be screened on a test criteria; for example, demographics or potential use of product. The questions asked of consumer panels should be answerable by untrained panelists.

© Lisa F. Young/ShutterStock, Inc.

FIGURE 3.2 Manufacturers, scientists, food technologists, and marketers can use consumer panels to gain a clear perception of what ordinary consumers may experience when tasting a particular food item.

Other Considerations

Other considerations also should be taken into account to optimize panelist performance during a sensory evaluation:

- It is wise to schedule the evaluation of certain product types at the time of day when that product is normally used or consumed.[5] For example, breakfast cereals would be better tested in the morning. In contrast, it would not be recommended to test highly flavored or alcoholic products in the early morning.
- Midmornings or midafternoons (such as 11 AM or 3 PM) are considered the best times for testing because at these times people are not usually overly hungry or full.[6]
- Panelists should not ingest any other food for at least 1 hour before testing and should not chew gum immediately before testing.[6]
- The instructions provided to the panelists should be very clear and concise. It is frequently desirable to give the instructions on how to perform the sensory evaluation verbally, before the panelists enter the booth area, and then also in written form on the score sheet.[2]
- Incentives usually are given as a token of appreciation to motivate people to participate voluntarily.[2] Common incentives include snacks or small gifts.

Environmental Controls

Physical and chemical factors present at the location of the sensory evaluation must be carefully controlled so that any possible extraneous effects of the surroundings on the test results are minimized and each panelist experiences the food in the same environment.

Temperature, Humidity, and Air Circulation

The ambient temperature should be comfortable, and the surroundings should be quiet and odor-free. The temperature and relative humidity for the sensory evaluation area should be 72–75°F (22–24°C) and 45–55%, respectively.[5] The use of replaceable active carbon filters in the ventilation system ducts is recommended. A slight positive pressure should be maintained in the booth areas to prevent odor contamination.[5] The sensory scientist should check if any unnecessary odors are detected in sensory testing areas.

Color and Lighting

The color and lighting in the sensory lab should be planned to permit adequate viewing of samples while minimizing distractions.[22,23] The walls of the sensory evaluation area should be off-white; the absence of hues of any color will prevent unwanted effects on food appearance.[5] Illumination in the booths should be uniform, shadow-free, and at least 300–500 lx at the table surface.[2] An ideal lighting system is controllable with a dimmer switch to a maximum of 700–800 lx, the common illumination intensity in offices.[2] Incandescent lights can control both the light intensity and the light color, but they generate heat, which will require cooling. Fluorescent lights generate less heat and allow for choice of whiteness; for example, cool white, warm white, or simulated north daylight.[5] Colored lights are used to mask visual differences among samples, calling for the subject to determine by flavor or texture only. A choice of low-intensity red, green, and/or blue lights using colored bulbs or special filters is a common feature of sensory booths.[5]

Product Controls

Variables pertaining to the product samples themselves must also be controlled.

Sample Preparation

Food samples must be of the same size (usually enough for two bites or sips) and from the same portion of the food (e.g., middle versus outside). The sensory analyst should determine and control the amount of product to be used in all the tests, including the amount of each added ingredient, the preparation process, and **holding time**, which is defined as the minimum and maximum time after preparation that a product can be used for a sensory test. For instance, suppose the test sample is a pumpkin muffin. The sensory analyst needs to decide the size of the muffin, exactly how the muffins will be baked, the appropriate holding time, and at which temperature and on what plate the sample will be served.

holding time The minimum and maximum time after preparation that a product can be used for a sensory test.

The sensory analyst should be very careful to standardize all serving procedures and sample preparation techniques except the variable under evaluation. If the appearance of the sample is not the variable under evaluation, then the samples should appear identical. Samples should be blind-labeled with random three-digit codes, and the sample order should be randomized to avoid bias due to order of presentation. A reasonable number of samples, say two to four, should be tested at a time to avoid taste fatigue.

Sample Temperature

Samples must be presented at the same temperature, which must be specified in the test protocol. For example, ice cream should be tempered at 5–9°F (–15°C to –13°C) for at least 12 hours before serving because scooping is not easy if the ice cream is colder.[2]

Presentation

Samples should be presented in containers or on plates that are the same size, shape, and color. White or clear containers are usually chosen so as not to influence panelists' perceptions of the food's color. The sensory analyst should choose the container that is most convenient. However, the choice of container should not negatively affect the flavor characteristics of the food product.

Carriers

Carriers 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 they are ingested, too.[2] Examples include spaghetti sauce on a spaghetti noodle, cream fillings in pastries, butter on bread, chips with salsa, and carrots with ranch dips. A carrier can mask or disguise differences or minimize the panelist's abilities to perceive the difference due to the addition of other flavors and modifications to texture and mouthfeel characteristics. However, for a product that is rarely consumed alone and almost always involves a carrier, the artificial situation where the carrier is not provided may affect test results, especially in consumer testing.[2] Therefore, whether a carrier is used should be carefully determined.

Palate Cleansers

Room temperature water or plain bread is made available for panelists to eat between samples to prevent carryover tastes. A rest period of at least 30 seconds is scheduled between samples. Paper towels or napkins are provided, and, because swallowing the food or beverage influences the taste of subsequent samples, small containers into which samples may be spit are provided.

One study that evaluated the effects of a range of palate cleansers (i.e., chocolate, pectin solution, table water crackers, warm water, water, whole milk) on foods representing various tastes and mouthfeels concluded that table water crackers were the only palate cleanser effective across all representative foods.[24] These foods included jelly beans (sweet), coffee (bitter), smoked sausage (fatty), tea (astringent), spicy tortilla chip (pungent), mint (cooling), and applesauce (nonlingering).

Measurement Theory

Measurement Hierarchy

Four levels of measurement are commonly used in sensory evaluation: nominal, ordinal, interval, and ratio.[25] It is important to recognize that there is a hierarchy in the level of measurement. At lower levels of measurement, assumptions tend to be less restrictive and data analyses tend to be less sensitive. At each level up the hierarchy, the current level includes all of the qualities of the one below it and adds something new. In general, it is desirable to have a higher level of measurement (such as interval or ratio) rather than a lower one (such as nominal or ordinal).[26]

Nominal

With nominal measurement, the numbers simply identify unique attributes; they are not ordered. For example, gender may be coded by assigning a "1" to males and a "2" to females. With nominal data, common descriptive statistics such as range, mean, or standard deviation are not appropriate. Instead, frequency counts, number of categories, or mode can be used to get some idea of the distribution of nominal data.

Ordinal

With ordinal measurement, the attributes can be ordered, but the difference between levels is not equal. It is not normally distributed and often is skewed. For example, cakes can be ordered by rank for perceived overall sweetness. The interval between values is not interpretable for ordinal measurement. In this case, the rank number can tell us where the cake falls in order of sweetness; however, we cannot draw conclusions about the differences among the products.

Interval

With interval measurement, there are ordered levels, and the difference between levels is equal. However, there is no true zero. For example, when we measure the oven temperature in degrees Fahrenheit, the distance from 100°F to 200°F is the same as the distance from 300°F to 400°F. Because the interval between the values is interpretable, it makes sense to calculate the average of the interval variable. However, in interval scaling, ratios do not make sense: 200°F is not twice as hot as 100°F, although the value is twice as large.

Ratio

With ratio measurement, there are ordered levels in which the difference between levels is equal, and there is a true zero. For example, weight is a ratio measurement. We can say that 200 pounds of sugar weighs twice as much as 100 pounds of sugar, and zero pounds of sugar means that there is no sugar.

Common Scales Used in Testing

Category Scales

Category scaling may be the oldest method of scaling; it involves the choice of discrete response alternatives to signify increasing sensation intensity in terms of degrees of liking and/or preference. The most popular category scale used in sensory testing is the **hedonic scale**, which measures the extent of like or dislike for the sensory characteristics of food. Examples of category scales are shown in **Figure 3.3**. Due to their simplicity, category scales are well suited for consumer panels. In addition, they offer some advantages in data coding and tabulation (for speed and accuracy) because they are easier to tabulate than line markings or the more variable magnitude estimates, described below.

hedonic scale A scale with which judges indicate the extent of their like or dislike for the sensory characteristics of food.

Line Scales

Line scales may also be referred to as graphic ratings or visual analog scales. Examples of line scales are shown in **Figure 3.4**. With line scales, the test participant's response is recorded as the distance of the mark from one end of the scale, usually whatever end is considered "lower." Line scaling differs from category scaling in the sense that the person's choices seem more continuous and less limited. Stone et al. recommended the use of line scaling for Quantitative Descriptive Analysis (QDA), then a relatively new approach to specifying the intensities of all the important sensory attributes.[27]

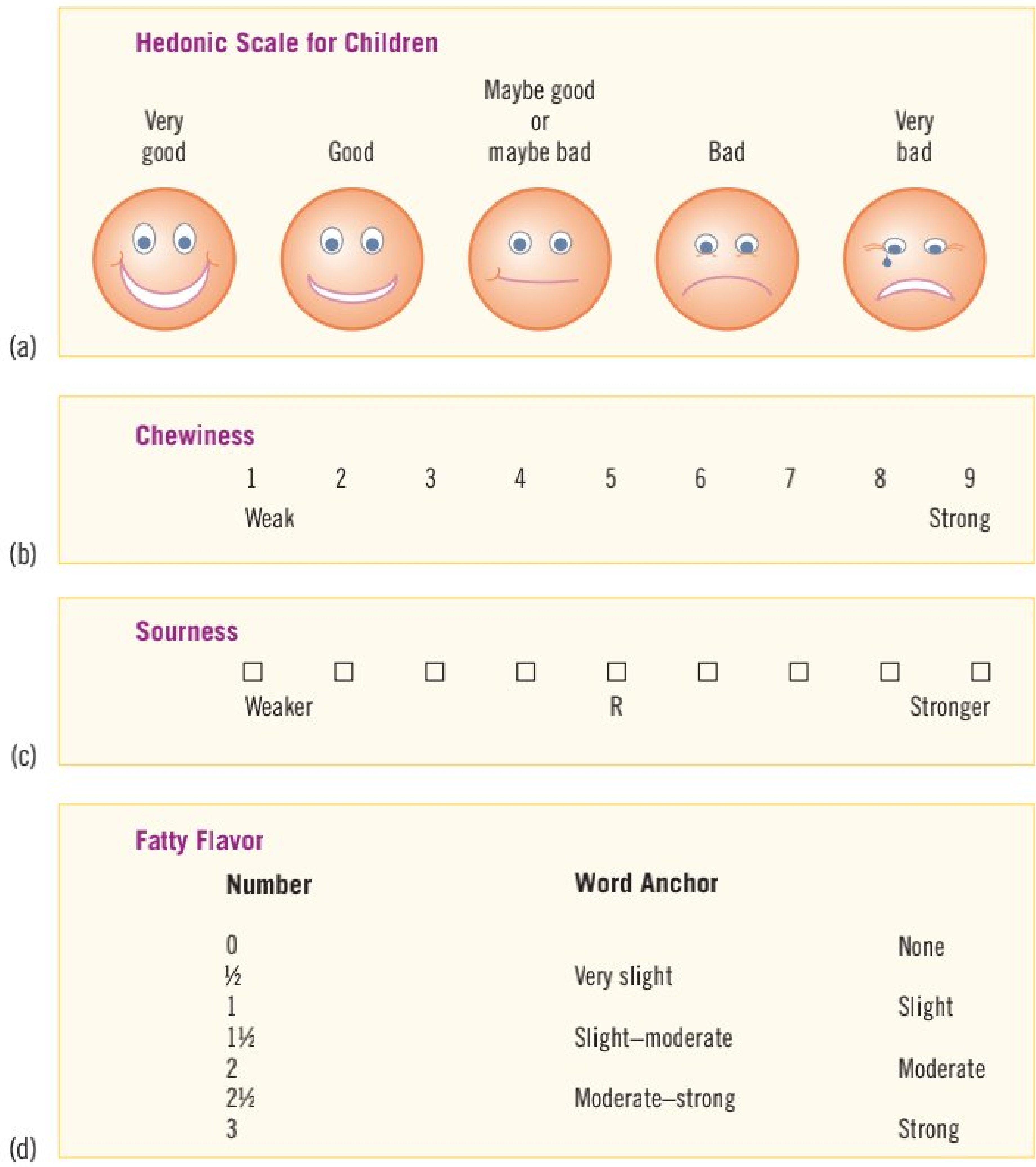


FIGURE 3.3 Examples of category scales. (a) Hedonic scale for children. (b) Chewiness. (c) Sourness. (d) Fatty flavor.

FIGURE 3.4 Examples of line scales. (a) Brownness. (b) Sweetness. (c) Roughness. (d) Overall acceptance.

Magnitude Estimation Scales

Magnitude estimation scaling is a popular technique in psychophysical studies. In this procedure, the panelists are instructed to assign numbers to their sensations in proportion to how strong each sensation feels.[2] In the analysis, the ratios between the numbers are supposed to reflect the ratios of sensation magnitudes that have been experienced. For example, if sample A is given the value of 15 for bitterness intensity and sample B seems three times as bitter, then B is given a magnitude estimation of 45.

Two variations of the magnitude estimation technique are available. In one method, a standard stimulus is given to the panelist as a reference with a fixed value, and all subsequent stimuli are rated relative to the reference. Values of zero are allowed in this method, but the rating of zero should not be used as a reference. In the second variation, no standard stimulus is given, and the panelist is free to choose any number for the first sample. All samples are then rated relative to the first intensity. For the second variation, because the panelists choose different ranges of numbers, the data have to be manipulated to bring all panelists into the same range. This adds an extra step to the analysis.

Panelists are cautioned to avoid falling into previous habits of using only category scales that they are used to using.[2] This may be a difficult problem with previously trained panels that have used a different scaling method because people like to stick with a method with which they are familiar.

Types of Sensory Tests

analytical tests Sensory tests used to detect discernible differences.

affective tests Sensory tests used to determine differences in acceptability or preference between products.

difference tests Sensory tests designed to detect discernible differences.

triangle test A difference test in which three samples are presented simultaneously (two of which are the same), and the judge is asked to identify the odd sample.

Sensory tests may be analytical or affective. **Analytical tests** are based on discernible differences, whereas affective tests are based on individual acceptability or preferences. Analytical tests are divided into two types of tests: difference (discriminative test) and descriptive. **Affective tests** have two categories, depending on the main task of the test: acceptance or preference. The primary task of an acceptance test is "rating," whereas the primary task of a preference task is "choice."[5]

Analytical Difference Tests

Difference tests are testing samples for their differences from each other. Difference tests can be used to test the sensitivity of judges as well as to perform a practical function such as determining whether a food company should buy an inexpensive ingredient to replace a more expensive one in formulating a food product. The two types of difference tests are overall difference tests and attribute difference tests.

Overall Difference Tests

The question of interest for an overall difference test is, "Does a sensory difference exist between samples?" Overall difference tests are the simplest sensory tests and include the triangle and the duo-trio test.

In a **triangle test**, three number-coded samples are presented simultaneously. The panelist is asked to indicate which one is odd (different from the other two). The difference in this method of presentation reduces the chance of guessing the right answer to 33.3% (1 in 3). This method is particularly effective in situations where treatment effects may have produced product changes that cannot be characterized simply by one or two attributes. It also can be useful for selecting panelists and monitoring their performance to discriminate given differences.

In a triangle test, six possible serving orders (AAB, ABA, BAA, BBA, BAB, ABB) are counterbalanced across all panelists. It should be conducted with 20 to 40 participants who have been screened for their sensory acuity to common product differences and who are familiar with the test procedure.

If differences are large and easy to identify, as few as 12 panelists may be employed.[5] The following is an example application of the triangle test:

- *Situation:* A food service management director wishes to confirm whether there is a significant sensory difference between a canned product and a paper carton product before changing the product. She is considering changing the chicken broth product from cans to cartons because of the clients' desires for "greener" packaging.
- *Test objective:* The objective is to determine if the change of product packaging causes an overall difference in sensory reactions to the chicken broth.
- *Test design:* The test design is shown in **Table 3.1**. There are 48 panelists with no replication.
- *Score sheet:* An example score sheet is shown in **Figure 3.5**.
- *Results and analysis:* The actual number of panelists who correctly identified the odd sample from the triangle test is counted. Out of 48 panelists, 25 correctly chose the odd sample. When a statistical test is applied, it is determined that the panelists could detect a difference between samples.
- *Conclusion*: A significant overall difference was found between the canned product and paper carton product.

Please taste the samples from left to right. Two samples are the same, and one is different. Circle the number of the sample that is different.

Sample code _______ _______ _______

Thank you!

FIGURE 3.5 Example of a triangle score sheet.

For detailed statistics regarding this example, see the Appendix at the end of this chapter.

The **duo-trio test** is another test of overall difference. In this test, the reference sample is presented first; it is followed by two other samples, one of which is the same as the reference. The panelists are requested to identify which of the latter two samples is the same as or different from the reference. With the duo-trio test, there is a 50% chance of being right by chance alone. The following example shows an application of the duo-trio test:

duo-trio test A difference test in which three samples are presented at the same time. A reference is designated, and the judge is asked to select the one most similar to the reference.

- *Situation:* A catering manager is faced with two very similar fluid egg products. Product A has been used for years; product B is a new product offered at a cheaper price. He wants to confirm whether there is an overall difference in perception between the two products.
- *Test objective:* The test objective is to determine whether the source of the fluid egg causes any overall difference in sensory perception of the scrambled egg.

TABLE 3.1
Triangle Test Design

Sample Identification	Codes in Sets with Two As	Codes in Sets with Two Bs
A: Canned product	624, 738	325*
B: Paper carton product	199*	801, 514
Panelist Number	**Serving Pattern**	**Codes in Order**
1, 7, 13, 19, 25, 31, 37, 43	AAB	624, 738, 199*
2, 8, 14, 20, 26, 32, 38, 44	ABA	624, 199*, 738
3, 9, 15, 21, 27, 33, 39, 45	BAA	199*, 624, 738
4, 10, 16, 22, 28, 34, 40, 46	BBA	801, 514, 325*
5, 11, 17, 23, 29, 35, 41, 47	BAB	801, 325*, 514
6, 12, 18, 24, 30, 36, 42, 48	ABB	325*, 801, 514

* Odd sample/correct answer.

- *Test design:* The test is conducted with 40 subjects who regularly eat eggs. Each of the two samples is used as the reference in half (20) of the evaluations. Scrambled eggs are made using the same method and equipment. The only variable that changes is the fluid egg product. Samples are presented without any condiment and at the same temperature.
- *Score sheet:* An example score sheet is shown in **Figure 3.6**.
- *Results and analysis:* Thirteen out of 40 panelists correctly matched the sample to the reference. Therefore, the panelists could not detect a difference between samples.
- *Conclusion:* The manager can conclude that there is no significant overall difference between the two fluid egg products.

For detailed statistics regarding this example, see the Appendix at the end of this chapter.

Please taste the samples from left to right. The left sample is a reference. Circle the number of the sample that matches the reference. If no difference is apparent between the two unknown samples, you must guess.

Reference _______
Sample code _______
Sample code _______

Thank you!

FIGURE 3.6 Example of a duo-trio score sheet.

Attribute Difference Tests

Attribute difference taste tests focus on a single sensory attribute such as sweetness or moistness. Attribute tests often are administered to evaluate qualitative differences in taste, color, and texture.

The **paired comparison test** is a test of difference in which a specific characteristic is designated. The subject is asked to test the two samples presented to identify the sample with the greater amount of the characteristic being measured. With this type of test, the subject has a 50% chance of being right by chance alone. The following example shows an application of the paired comparison test:

paired comparison test
A difference test in which two samples are presented, and the judge is asked to select the one that has more of a particular characteristic.

ranking test A difference or preference test in which more than two samples are presented and all samples are compared by ranking them from lowest to highest for the intensity of a specific characteristic.

- *Situation:* A cookie manufacturer receives reports from the market that her cookie (cookie A) is deemed insufficiently sweet, so a test cookie (cookie B) is made using more sweetener. She wants to produce a cookie that is perceptibly sweeter, but not excessively so.
- *Test objective:* The test objective is to compare cookie A with cookie B to determine whether a small but significant increase in sweetness has been attained.
- *Test design:* A paired comparison test is chosen because the characteristic of interest is only sweetness, nothing else. The sensory analyst codes the cookies "796" and "308" and offers them to a panel of 30 subjects with proven ability to detect small changes in sweetness. The panelists are not asked, "Is 308 sweeter than 796?" but rather, "Which cookie is sweeter?" so as to eliminate the potential for bias.
- *Score sheet:* An example score sheet is shown in **Figure 3.7**.
- *Results and analysis:* Significantly more panelists (22) identified test cookie B as being sweeter.
- *Conclusion:* There is a significant sweetness difference between cookies A and B. The test cookie was successful.

There are two samples in each of the two paired comparison sets for you to evaluate. Taste each of the coded samples in the sequence presented from left to right. Please write down the code of the sweeter sample.

Sample code _______ Sample code _______

Which sample is sweeter? _______

Thank you!

FIGURE 3.7 Example of a paired comparison score sheet.

For detailed statistics regarding this example, see the Appendix at the end of this chapter.

The **ranking test** is valuable when several samples need to be evaluated for a single characteristic. Ranking test procedures have the advantage of simplicity in instructions to panelists, ease of data handling, and minimal assumptions about the level of measurement because the data are ordinal.

Although ranking tests are most often applied to hedonic data, they also are applicable to questions of sensory intensities. The following example shows an application of the ranking test:

- *Situation:* A sandwich maker wishes to compare the saltiness among different types of cheese to select a cheese for a new sandwich menu.
- *Test objective:* The test objective is to determine whether there is a significant difference in saltiness among three types of cheese (mozzarella, cheddar, and provolone).
- *Test design:* The ranking test is suitable because it is simple to carry out and does not require much training. The three samples are tested with a panel of 28 students. The panelists receive the samples (coded with three-digit numbers) in balanced, random order.
- *Score sheet:* An example score sheet is shown in **Figure 3.8**.
- *Results and analysis:* When compared to the minimum critical differences, no significant differences are found among the samples because all of the differences are less than 18.
- *Conclusion:* There is no significant difference in saltiness among the three types of cheese.

For detailed statistics regarding this example, see the Appendix at the end of this chapter.

The **rating difference test** among multiple samples is used when a rating scale is applied. The following example shows an application of the rating difference test:

- *Situation:* A baker is producing a new low-fat pound cake prepared by replacing butter with applesauce. He is making four samples with four levels of applesauce substitution (0%, 33%, 66%, and 100%). He discovers that moistness is the sensory characteristic that shows the most noticeable difference among samples; thus, it is the variable most likely to affect product acceptability.
- *Test objective:* The test objective is to compare the moistness of four pound cakes in order to determine which product has no significant difference in moistness from the control product and has the lowest fat content.
- *Test design:* A completely randomized design is used, in which all samples are compared together using a rating scale. Twelve subjects evaluate the moistness for four samples on scale of 1 to 9.
- *Score sheet:* An example score sheet is shown in **Figure 3.9**.
- *Results and analysis:* There are significant differences in moistness between 100% applesauce pound cake and the other three products (control with butter and no applesauce, and 33% applesauce, and 66% applesauce).
- *Conclusion:* The 66% applesauce pound cake should be chosen because it is the product with the lowest fat content that shows no significant difference from the control in moistness.

For detailed statistics regarding this example, see the Appendix at the end of this chapter.

Please taste each of the coded samples in the set in the sequence presented, from left to right. Rank the three samples in descending order of saltiness. You may re-taste any of the samples while ranking for intensity of the saltiness. No ties are allowed in the ranking. Rinse your mouth with water between samples, and wait for 30 seconds before you taste the next sample. Remember that the most intense sample should be ranked 1.

Saltiness

Sample code ______ ______ ______
1 2 3

Thank you!

FIGURE 3.8 Example of a ranking score sheet.

rating difference test Test to differentiate among multiple samples that uses a rating scale. Products are ranked using a rating scale to assess for differences between samples.

Sample code ______

Moistness
1 2 3 4 5 6 7 8 9
Weakest Strongest

Thank you!

FIGURE 3.9 Example of a rating difference score sheet.

descriptive tests Sensory tests designed to provide information on the specific sensory characteristics of food samples and to quantify the sensory differences.

profiling A group of highly trained panelists work together to develop the vocabulary needed to provide specific descriptions of food samples; used to detail the specific flavors (flavor profiling) or textures (texture profiling) of a food or beverage.

Innovative Point

Approaches to Descriptive Tests The four major approaches of descriptive tests are flavor profile, quantitative descriptive analysis, texture profile, and sensory spectrum.

Analytical Descriptive Tests

The descriptive sensory test is the most comprehensive and informative test used in sensory evaluation. The question for the descriptive test is, "How do products differ in specific sensory characteristics?" **Descriptive tests** enable researchers to characterize their products through selective, critical scoring of specific attributes of each product. The descriptive techniques are frequently used for developing new products and for quality assurance. The information from these methods can be especially valuable for dealing with sensory problems that consumers may detect.

Descriptive tests require a well-trained panel and tend to be expensive. Descriptive tests should never be used with consumers because consistent and reproducible data are an essential part of descriptive tests.[2] Descriptive testing is usually conducted by using a scorecard containing precise word descriptions that structure the form of the responses. Each of the characteristics of a sample to be evaluated is described over a range, and the panelist selects the specific description matching the sample for each item on the scorecard. The responsibility for selecting the appropriate vocabulary to elicit an accurate picture of the samples rests with the researcher who has developed the scorecard; a well-constructed scorecard will give the desired information.[2]

Profiling is another approach to descriptive testing. To conduct **profiling**, a group of highly trained panelists work together to develop the vocabulary needed to provide specific descriptions of food samples. This technique is used to detail the specific flavors (flavor profiling) or textures (texture profiling) of a food or beverage.

Flavor Profile Analysis®

Flavor Profile Analysis (FPA) is based on the concept that flavor consists of identifiable taste, odor, and chemical-feel factors as well as an underlying complex of sensory impressions that are not separately identifiable.[28] Scientists at Arthur D. Little developed this technique in the late 1940s and early 1950s, and the name and the technique were trademarked to Arthur D. Little and Co.[2] The method is a qualitative descriptive test that involves formal procedures for describing and assessing the aroma, flavor, and aftertaste of a product in a reproducible manner. FPA is a consensus technique.[2] The vocabulary used to describe the sample and the sample evaluation itself are achieved by panel members who work together to come to an agreement.

Quantitative Descriptive Analysis®

Quantitative Descriptive Analysis (QDA) was developed during the 1970s to correct some of the perceived problems associated with FPA.[1,27] Unlike FPA data, QDA data are not generated through consensus discussion, and panel leaders are not active participants. Unstructured line scales are used to describe the intensity of rated sensory attributes. Stone et al.[27] chose the linear graphic scale, a line that extends beyond fixed verbal endpoints, based, in part, on earlier studies.[29]

In QDA, 10 to 12 panelists begin their training by generating a consensus vocabulary.[27] They are exposed to many possible variations of the product to facilitate the acquisition of an accurate concept, and the panel leader acts only as a facilitator by directing discussion and supplying materials such as references and other samples required by the panel.[2] The actual product evaluations are performed by each panelist individually, usually in an isolation booth.

Texture Profile Analysis®

Texture profile analysis (TPA) was developed by scientists working for General Foods during the 1960s and was subsequently modified by several sensory specialists.[30–35] The TPA uses a standardized terminology to describe textural characteristics by both their physical and sensory aspects.[2] Definitions and listing order of the terms are determined through consensus by the TPA panelists. The reference scales anchor both the range and the concept for each term.[35] The full range of a specific parameter by reference products helps panelists confirm the intensity increments within each scale. The use of the same reference frame is the key to a successful TPA. Sample preparation, presentation, and evaluation should be strictly controlled. Panelists should also be trained to bite, chew, and swallow in a standardized way.

Sensory Spectrum®

Gail Civille, who became a TPA expert at General Foods, subsequently created the Sensory Spectrum (SS) technique.[2] This technique is an expansion on descriptive analysis techniques. The unique characteristic of the SS technique is that panelists do not generate a panel-specific vocabulary to describe sensory attributes of products, but rather use a standardized word list (lexicon).[36] The terminology used to describe a particular product is chosen *a priori* and remains the same for all products within a category over time.[2] Because panelists are trained to use the scales in an identical manner, use of the SS technique should allow data from experiments that include only one sample to be compared against data from different samples used in other studies.

Panelist training for the SS technique is much more extensive than that for QDA, and the panel leader has a more directive role than in QDA.[2] Similar to the TPA, panelists are provided lexicons that are used to describe perceived sensations associated with the samples. The panelists use a numerical intensity scale—usually a 15-point scale—and they are supplied with reference standards.

Affective Tests

For a food product to be successful in the marketplace, consumers must prefer it over other products. Therefore, consumer panels often are used to indicate preference of one sample over another. The panelist rates his or her preference for one of the samples on a specific quality on the score sheet. Hedonic rating scales can be used to measure the degree of pleasure experienced with each sample. Sometimes, the frequency that a panelist might desire to eat the sample is measured as a way to determine the acceptability of the various samples.

Acceptance Tests

Acceptance tests involve rating the difference in acceptance between two samples. The following example shows an application of a rating acceptance test:

- *Situation:* A restaurant manager wishes to compare acceptance between two types of chocolate candies (products A and B). She wants to give the more-liked chocolate product as a customer appreciation present at the end of a meal.
- *Test objective:* The test objective is to determine which chocolate product is more liked.
- *Test design:* 20 panelists are asked to evaluate acceptance for the two chocolate products using 5-point hedonic scales. Two samples (coded with random three-digit numbers) are presented simultaneously at the same temperature. Half of the panelists test product A first; the other half test product B first.
- *Score sheet:* An example score sheet is shown in **Figure 3.10**.

Please taste each sample, and indicate how well you like it.

318 (A)			442 (B)		
______	Like very much	(5)	______	Like very much	(5)
______	Like moderately	(4)	______	Like moderately	(4)
______	Neutral	(3)	______	Neutral	(3)
______	Dislike moderately	(2)	______	Dislike moderately	(2)
______	Dislike very much	(1)	______	Dislike very much	(1)

Thank you!

FIGURE 3.10 Example of a 5-point acceptance score sheet.

- *Results and data analysis:* Product A has a higher acceptance rate than product B. The calculated *t*-value for the difference values exceeds the reference *t*-value and is statistically significant.
- *Conclusion:* Product A is selected as a consumer appreciation present because product A is significantly more acceptable than product B.

For detailed statistics regarding this example, see the Appendix at the end of this chapter.

Sometimes, rating acceptance tests involve more than two samples. The following is an example application of a rating acceptance test that uses multiple samples:

- *Situation:* A dietitian is developing gluten-free chocolate chip cookies prepared by replacing all-purpose flour with brown rice and chickpea flours.
- *Test objective:* The test objective is to determine whether the gluten-free chocolate chip cookies (100% brown rice flour, 50% brown rice/50% chickpea flour, and 100% chickpea flour) are sufficiently acceptable against 100% all-purpose flour cookies.
- *Test design:* A completely randomized design that compares all samples together using a rating scale is used. One hundred subjects evaluate the overall acceptability of four samples on a 9-point hedonic scale.
- *Score sheet:* An example score sheet is shown in **Figure 3.11**.
- *Results and analysis:* The 100% chickpea flour cookie product has a significantly higher acceptance rate than other gluten-free cookies (50% brown rice/50% chickpea flour cookies and 100% brown rice flour cookie products); however, there is no significant difference among the control (all-purpose flour) and 100% chickpea flour cookies.
- *Conclusion:* The 100% chickpea flour cookies are acceptable gluten-free alternatives to the conventional all-purpose flour cookie.

For detailed statistics regarding this example, see the Appendix at the end of this chapter.

Taste each sample, and indicate how well you like it.

Sample code ______

☐ Like extremely
☐ Like very much
☐ Like moderately
☐ Like slightly
☐ Neither like nor dislike
☐ Dislike slightly
☐ Dislike moderately
☐ Dislike very much
☐ Dislike extremely

Thank you!

FIGURE 3.11 Example of a 9-point verbal hedonic scale sensory sheet.

Preference Tests

The question of interest for the preference test is, "Which sample do you prefer?" The following example shows an application of a paired preference test:

- *Situation:* A caterer wants to know determine which brand of soft drink she should use. She wants to compare the preferences for two prospective products.
- *Test objective:* The test objective is to determine which soft drink product is preferred over the other product.
- *Test design:* One hundred subjects who are soft-drink drinkers are invited to a central location where the company caters. Two products (A and B) coded with three-digit random numbers are presented simultaneously at the same temperature. Half the participants receive the soft drinks in the order A–B, and the other half receive them in the order of B–A. The serving temperature and serving time after opening the containers are carefully controlled. (Both are particularly important variables to control for carbonated beverages.)
- *Score sheet:* An example score sheet is shown in **Figure 3.12**.

Taste the sample on the left first and the sample on the right second. Which one do you prefer? Please choose one.

337 ☐ 198 ☐

Thank you!

FIGURE 3.12 Example of paired preference score sheet.

- *Results and analysis:* The number (66) that chose sample A is larger than the critical value (61) at a level of statistical significance.
- *Conclusion:* There is a significant preference difference between soft drink products. The analyst recommends that the caterer serve soft drink product A.

For detailed statistics regarding this example, see the Appendix at the end of this chapter.

The final example in this chapter details the application of a ranking preference test:

- *Situation:* A school dietitian wishes to compare preferences for three different varieties of apples.
- *Test objective:* The test objective is to determine whether there is a significant difference in preference among three types of apples (products A, B, and C).
- *Test design:* The three samples are ranked by 100 panelists. Each panelist receives three samples coded with three-digit numbers and served in balanced, random order.
- *Score sheet:* An example score sheet is shown in **Figure 3.13**.
- *Results and analysis:* When compared to the minimum critical differences at $\alpha = 0.05$ (34), product B is preferred over products A and C. There is no significant preference difference between products A and C.
- *Conclusion:* It is recommended that the school serve apple B, which is preferred over apples A and C.

For detailed statistics regarding this example, see the Appendix at the end of this chapter.

Please taste each of the coded samples in the set in the sequence presented from left to right. Rank the three samples in descending order of preference. You may re-taste any of the samples while ranking for the preference. No ties are allowed in the ranking. Remember that the most preferred sample should be ranked 1.

Sample code ______ ______ ______
1 2 3

Thank you!

FIGURE 3.13 Example of a preference ranking score sheet.

Chapter Review

Sensory evaluation is a scientific testing method for accurate measurement of human responses as perceived by the five senses. Sensory evaluation is a vital part of food development because it is the essential means of determining how consumers will react to a food. Reliable sensory evaluation can be performed by optimizing four steps: definition of the problem, test design, instrumentation, and interpretation. People evaluate a particular food primarily based on how it looks, smells, tastes, sounds, and feels. The food attributes that are typically perceived through the human senses are appearance, odor, taste, flavor, consistency, and texture.

The environment in which the sensory test is conducted should be carefully controlled, and samples must be prepared and presented in a uniform fashion so as not to influence panelists' perception of the food's quality. Panelists who are well suited to the purpose of the sensory test should be selected and trained appropriately.

The two types of sensory tests are analytical and affective. Analytical tests are based on discernible differences, whereas affective tests are based on individual acceptability or preferences. Analytical tests are divided into two types of tests: difference tests (discriminative tests) and descriptive tests. Depending on the main task of the test, affective tests are either acceptance tests or preference tests. The primary task of acceptance tests is to rate the degree of liking, whereas with preference tests the goal is to identify the item that is more liked.

Learning Portfolio

Key Terms

Study Points

1. Sensory evaluation is a scientific testing method for accurate measurement of human responses as perceived by the five senses.
2. The food attributes that are typically perceived through the five senses are appearance, odor, taste, flavor, consistency, and texture.
3. The environment in which the sensory test is conducted, as well as sample preparation and presentation, should be carefully controlled so as not to bias the results of the test. Panelists who are well suited to the purpose of the sensory test should be selected and trained appropriately.
4. The two types of sensory tests are analytical and affective. Analytical tests are based on discernible differences, whereas affective tests are based on acceptability or preferences.
5. Analytical tests include difference tests (discriminative tests) and descriptive tests. The two types of difference tests are overall difference tests and attribute difference tests.
6. The triangle test is an overall difference test. In the triangle test, three samples are presented simultaneously; two samples are alike and one is different. Panelists are asked to indicate the odd sample. The chance of obtaining a correct answer by guessing is 33.3% (1 in 3) in the triangle test.
7. The duo-trio test is an overall difference test. In the duo-trio test, the reference sample is presented first; it is then followed by two other samples, one of which is the same as the reference. The judge is asked to identify which of the last two samples is same as or different from the reference. There is a 50% chance of being right by simply guessing in a duo-trio test.
8. The paired comparison test is an attribute difference test in which a specific characteristic is designated. The subject is asked to test the two samples presented to identify the sample with the greater amount of the characteristic being measured. The subject has a 50% chance of being right by chance alone.
9. Descriptive sensory tests are the most comprehensive and informative tools used in sensory evaluation. Descriptive tests require a well-trained panel and should never be used with consumers because consistent and reproducible data are an essential part of descriptive tests. Examples of descriptive tests include the Flavor Profile Analysis (FPA), Quantitative Descriptive Analysis (QDA), Texture Profile Analysis (TPA), and the Sensory Spectrum (SS) technique.
10. Affective tests include acceptance tests and preference tests. The primary task of acceptance testing is to rate the degree of liking, whereas the preference task seeks to identify the product that is liked more.

Issues for Discussion

1. Are sensory preferences due to genetic or environmental causes?
2. Should the food industry provide foods with preferred tastes that may be sugar and salt based or try to provide healthier foods, knowing that the taste sensations may not be what most consumers prefer? Is there an ethical choice?

Research Ideas for Students

1. Taste preferences of different ethnic groups and possible reasons for such differences
2. Taste preferences of cancer patients

References

1. Stone H, Sidel JL. *Sensory Evaluation Practices.* 3rd ed. San Diego, CA: Academic Press; 2004.
2. Lawless HT, Heymann H. *Sensory Evaluation of Food: Principles and Practices.* 2nd ed. New York: Springer; 2010.

3. Meiselman HL. Critical evaluation of sensory techniques. *Food Qual Pref.* 1993;4:33–40.
4. Pfenninger HB. Methods of quality control in brewing. *Schweizer Brauerei-Rundschau*. 1979;90:121.
5. Meilgaard MC, Civille GV, Carr BT. *Sensory Evaluation Techniques.* 4th ed. Boca Raton, FL: CRC Press; 2007.
6. Brown A. *Understanding Food: Principles and Preparation.* 3rd ed. Belmont, CA: Thompson-Wadsworth; 2008.
7. Maruniak JA. The sense of smell. In: Piggott JR (ed.). *Sensory Analysis of Foods.* 2nd ed. London: Elsevier; 1988.
8. McWilliams M. *Foods: Experimental Perspectives*. 6th ed. Upper Saddle River, NJ: Pearson/Prentice-Hall; 2008.
9. Kim UK, Jorgenson E, Coon H, et al. Positional cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. *Science*. 2003;299(5610):1221–1225.
10. Blakeslee AF. Genetic of sensory thresholds: taste for phenylthiocarbamide. *ProcNatlAcadSci USA.* 1932:120–130.
11. Fox AL. The relationship between chemical constitution and taste. *Proc Natl Acad Sci USA.* 1932:115–120.
12. Lawless HT. A comparison of different methods for assessing sensitivity to the taste of phenylthiocarbamide PTC. *Chem Senses.* 1980;5:247–256.
13. Bartoshuk LM, Duffy VB, Miller IJ. PTC/PROP tasting: anatomy, psychophysics, and sex effects. *Physiol Behav.* 1994;56(6): 1165–1171.
14. Guo SW, Reed DR. The genetics of phenylthiocarbamide perception. *Ann Hum Biol.* 2001;28(2):111–142.
15. Tepper B, Nurse R. Fat perception is related to PROP taster status. *Physiol Behav*. 1997;61(6):949–954.
16. Reedy F Jr, Bartoshuk LM, Miller I, et al. Relationship among papillae, taste pores, and 6-n-propylthiouracil (PROP) suprathreshold taste sensitivity. [abstract]. *Chem Senses*. 1996;21:616.
17. Hosako NY, Lucchina LA, Synder DJ, et al. Number of fungiform papillae in non-tasters, medium tasters and supertasters of PROP (6-n-propylthiouracil). [abstract]. *Chem Senses.* 1996;21:616.
18. Duffy VB, Bartoshuk LM. Food acceptance and genetic variation in taste. *J Am DietAssoc.* 2000;100:647–655.
19. Duffy VB, Peterson JM, Dinehart ME, Bartoshuk LM. Genetic and environmental variation in taste. *Top Clin Nutr*. 2003;18(4): 209–220.
20. Godshall MA. How carbohydrates influence food flavor. *Food Tech.* 1997;51(1):62–67.
21. De Roos KB. How lipids influence food flavor. *Food Tech.* 1997;51(5):60–62.
22. Amerine MA, Pangborn RM, Roessler EB. *Principles of Sensory Evaluation of Food.* New York: Academic Press; 1965.
23. International Organization for Standardization (ISO). Sensory analysis-general guidance for the design of test rooms. In: International Standard ISO 8589. Geneva, Switzerland: International Organization for Standardization; 1998.
24. Lucak CL, Delwiche JF. Efficacy of various palate cleansers with representative foods. *Chemosensory Perception*. 2009;2:32–39.
25. Stevens SS. Mathematics, measurement and psychophysics. In: Stevens SS (ed.). *Handbook of Experimental Psychology*. New York: Wiley; 1951: 1–49.
26. Trochim WM. *Research Methods: The Concise Knowledge Base*. Cincinnati, OH: Atomic Dog Pub; 2005.
27. Stone H, Sidel J, Oliver S, et al. Sensory evaluation by quantitative descriptive analysis. *Food Tech.* 1974;28:24–29, 32, 34.
28. Hootman RC (ed.). *Manual on Descriptive Test for Sensory Evaluation*. ASTM Manual Series MNL 13. Philadelphia, PA: American Society for Testing and Materials; 1992.
29. Anderson NH. Functional measurement and psychological judgment. *Psych Rev.* 1970;153–170.
30. Brandt MA, Skinner EZ, Coleman JA. The texture profile method. *J Food Sci.* 1963;28:404–409.
31. Civille GV, Liska IH. Modifications and applications to foods of the general foods sensory texture profile technique. *J Texture Studies.* 1975;6:19–31.
32. Szczesniak AS. Texture measurements. *Food Tech.* 1966;20: 1292–1298.
33. Szczesniak AS. General foods texture profile revisited—Ten years perspective. *J Texture Studies*. 1975; 6:5–17.
34. Szczesniak AS. Classifications of textural characteristics. *J Food Sci.* 1963;28:385–389.
35. Szczesniak AS, Brandt MA, Friedman HH. Development of standard rating scales for mechanical parameters of texture and correlation between the objective and the sensory methods of texture evaluation. *J Food Sci.* 1963;28:397–403.
36. Civille GV, Lyon B. *ASTM Lexicon Vocabulary for Descriptive Analysis*. Philadelphia, PA: American Society for Testing and Materials; 1996.

Appendix

Sensory Evaluation

Data Analysis for the Triangle Test

Step 1: Refer to the reference minimum number of correct responses for the triangle test.[1]

- *Application*: Because the number of panelists (trials) in this example is 48, the minimum number of correct responses corresponding to the probability level of 0.05 is 22.[1]

Step 2: Count the number of panelists who correctly identified the odd sample.

- *Application*: Out of 48 panelists, 25 correctly chose the odd sample.

Step 3: If the value (step 2) is same or larger than the reference value (step 1), we can say that the panelists could detect a significant difference between the samples at a probability of 5%.

- *Application*: The value (25) from step 2 is larger than the reference value. Therefore, the panelists could detect a difference between samples.

Data Analysis for the Duo-Trio Test

Step 1: Refer to the critical number of correct responses in the duo-trio test.[1]

- *Application*: Because the number of panelists (trials) in this example is 40, the critical number of correct responses corresponding to the probability level of 0.05 is 26.[1]

Step 2: Count the number of panelists who correctly matched the sample to the reference.

- *Application*: Out of 40 panelists, 13 correctly matched the sample to the reference.

Step 3: If the value (step 2) is the same or larger than the table value (step 1), we can say that the panelists could detect a significant difference between the samples at a probability of 5%.

- *Application*: The value (13) from step 2 is smaller than the reference value. Therefore, the panelists could not detect a difference between samples.

Data Analysis for the Paired Comparison Test

Step 1: Refer to the reference minimum number of correct judgments for the paired comparison test.[1]

- *Application*: In our case, the number of trials is 28 and the probability level is 0.05. The reference value is 19.[1]

Step 2: Count the number of panelists who chose sample A or B, respectively.

- *Application*: Six panelists chose sample 796 (cookie A, the original) as the sweeter sample and 22 panelists chose sample 308 (cookie B, the new one).

Step 3: If the value (step 2) is the same or larger than the reference value (step 1), we can say that the sample that was chosen more often is significantly different (sweeter) than the other sample.

- *Application*: The value (22) from step 2 is larger than the table value (19). Therefore, the panelists could detect a difference in sweetness between samples.

Data Analysis for the Ranking Test

The ranking data can be analyzed either by using the Basker's tables[2] or those by Newell and MacFarlane.[3] This analysis used a reworked table[1] from Newell and MacFarlane's.[3] **Table A** shows how the results are organized. The ranks are summed, and the differences between sums are compared to the critical values in the table. From the table in the reference, the corresponding number of samples on the horizontal axis (in our example, 3 samples) and the number of panelists from the vertical axis (in our example, 28 panelists) are selected. The number where the two points cross will be the critical values for the difference. In our case, the critical value is 18. If the difference of sum of their rank between each pair of samples is greater than the critical value, it is considered that the samples are significantly different in saltiness.

TABLE A
Results of the Ranking Test for the Saltiness of Three Types of Cheeses

Panelist No.	*A* (183, Mozzarella)	*B* (479, Cheddar)	*C* (862, Provolone)
1	3	2	1
2	3	2	1
3	2	3	1
4	2	1	3
5	1	2	3
6	3	1	2
7	3	2	1
8	2	1	3
9	2	1	3
10	2	3	1
11	2	1	3
12	2	1	3
13	2	1	3
14	1	3	2
15	2	1	3
16	2	1	3
17	3	2	1
18	3	2	1
19	2	1	3
20	2	3	1
21	3	1	2

(*continues*)

TABLE A
Results of the Ranking Test for the Saltiness of Three Types of Cheeses (*continued*)

Panelist No.	*A* (183, Mozzarella)	*B* (479, Cheddar)	*C* (862, Provolone)
22	1	2	3
23	1	2	3
24	3	1	2
25	1	2	3
26	3	1	2
27	3	1	2
28	3	2	1
Column sum	62	46	60
Differences	*A* vs. *B* = 16	*B* vs. *C* = 14	*A* vs. *C* = 2

Data Analysis for the Rating Difference Test

The data is evaluated by one-way analysis of variance (ANOVA). There are 4 treatments and 12 observations per treatment (see **Table B**). Because the *F*-value (26.65) is very significant ($P < 0.0001$) (see **Table C**), there are significant differences between treatments.

TABLE B
Results of the Rating Test for the Moistness of Four Types of Pound Cake

Panelist No.	Treatment			
	0% Applesauce (Control)	33% Applesauce	66% Applesauce	100% Applesauce
1	9	8	8	5
2	8	8	7	6
3	9	7	7	4
4	6	5	5	3
5	7	8	6	5
6	8	9	8	5
7	9	7	7	5
8	9	8	8	4
9	8	7	8	4
10	7	8	7	5
11	7	6	6	5
12	7	8	8	4

TABLE C
Completely Randomized Design One-Way Analysis of Variance of Results in Table B

Source of Variation	Sum of Squares	Degree of Freedom	Mean Square	*F*-value	*P*-value
Treatments	77.06	3	25.69	26.65	< 0.0001
Error	42.42	44	0.96		
Total	119.48	47			

To determine which samples are significantly different, perform a Tukey Honestly Significant Difference (HSD) multiple comparison test. The results are as follows:

Control, 0% applesauce	7.83b
33% applesauce	7.42b
66% applesauce	7.08b
100% applesauce	4.58a

Data Analysis for the Affective Test: Chocolate Candies

The result of the acceptance test for two types of chocolate candies and the *t*-value calculation are shown in **Table D**. In this case, the critical *t*-value (degree of freedom = 19) from the Student's *t*-distribution table at $\alpha = 0.05$ is 2.093 (a student's *t*-distribution table can be found in most statistics books or online [see http://www.itl.nist.gov/div898/handbook/eda/section3/eda3672.htm]). Because the calculated *t*-value (2.62) is larger than the reference *t*-value, a significant acceptance difference exists between the samples.

TABLE D
Results of the Acceptance Test for Two Types of Chocolate Candies

Panelist	Original (*A*)	Modified (*B*)	Difference ($A - B = D$)	D^2
1	4	4	0	0
2	5	3	2	4
3	3	3	0	0
4	4	2	2	4
5	4	3	1	1
6	4	3	1	1
7	3	4	–1	1
8	4	4	0	0
9	5	3	2	4
10	4	4	0	0
11	3	3	0	0
12	5	4	1	1
13	4	2	2	4
14	3	3	0	0
15	4	4	0	0
16	5	4	1	1
17	4	5	–1	1
18	5	4	1	1
19	3	3	0	0
20	5	5	0	0
Sum	81	70	11	23

Sum of $D = 11$, Mean of $D = 0.55$, Sum of $D^2 = \sum D^2 = 23$

$$\sqrt{\frac{\sum D^2 - (\sum D)^2/N}{N - 1}}$$

Standard deviation (SD) of $D =$

$$\sqrt{\frac{23 - (121/20)}{20 - 1}} = \sqrt{\frac{23 - 6.05}{19}} = \sqrt{\frac{16.95}{19}} = \sqrt{0.892} = 0.94$$

Standard error (SE) of D = SD of $D/\sqrt{N} = 0.94/\sqrt{20} = 0.94/4.47 = 0.21$
t = Mean of D/SE of $D = 0.55/0.21 = 2.62$

Data Analysis for the Affective Test: Cookie Formulations

The result of one-way analysis of variance for the overall acceptance of gluten-free cookies is shown in **Table E**.

TABLE E
Completely Randomized Design One-Way Analysis of Variance for the Overall Acceptance of Gluten-Free Cookies

Source of Variation	Sum of Squares	Degree of Freedom	Mean Square	*F*-value	π-value
Treatments	968.84	3	322.95	439.23	< 0.0001
Error	291.16	396	0.74		
Total	1260.00	399			

Because the *F*-value is statistically significant, there are significant differences among treatments. Therefore, the Tukey Honestly Significant Difference (HSD) Post-hoc test was performed. The results of the Tukey HSD test are as follows:

Control, 100% all-purpose flour cookies	7.87c
100% brown rice flour cookies	4.21a
50% brown rice/50% chickpea flour cookies	6.57b
100% chickpea flour cookies	8.15c

Data Analysis from the Paired Preference Test

Step 1: Refer to the reference minimum number of agreeing judgments for the paired preference test.[2]

- *Application:* In our case, the number of trials is 100, and the probability level is 0.05. The critical reference value is 61.[1]

Step 2: Count the number of subjects who chose sample *A* or *B*, respectively.

- *Application*: 66 subjects chose sample 337 (soft drink *A*) as the preferred sample and 34 panelists chose sample 108 (soft drink *B*).

Step 3: If the value (step 2) is same or larger than the critical value (step 1), we can say the sample that was chosen more is significantly preferred over the other sample.

Data Analysis for Ranking Preference Test

Ranks are added, and the differences between the sums are compared to the critical reference value.[1] Similar to the data analysis for the difference ranking test, select the corresponding number of samples on the horizontal axis and then select the number of panelists from the vertical axis in the reference table.[1] Find the number where the two points cross. In this example, because 100 panelists ranked three products, the critical value is 34.[1] If the difference of the rank sum between each pair of samples is greater than the critical value, then the samples are significantly different in preference. Because the rank scale used was 1 = preferred most and 3 = preferred least, the smallest rank sum means that that the product is the most preferred.

In **Table F**, because the differences of rank sum between product *B* and the other products (*A* and *B*, *B* and *C*) are larger than the critical value (34), product *B* is significantly preferred over the other products.

TABLE F
Results of the Ranking Preference Test for Three Varieties of Apples

	908 (*A*)	144 (*B*)	862 (*C*)
Rank sum	216b	164a	220b
Differences	*A* vs. *B* = 52	*B* vs. *C* = 56	*A* vs. *C* = 4

References

1. Lawless HT, Heymann H. *Sensory Evaluation of Food: Principles and Practices.* 2nd ed. New York: Springer; 2010: 563, 565, 566.
2. Basker D. Critical values of differences among rank sums for multiple comparisons. *Food Tech.* 1988 July:88–89.
3. Newell GJ, MacFarlane JD. Expanded tables for multiple comparison procedures in the analysis of ranked data. *J Food Sci.* 1987;52:1721–1725.