Sensory Evaluation: Analytical and Affective Testing

By Marianne Gillette, McCormick & Company, Inc., Hunt Valley, Maryland

Probably the most critical item in sensory evaluation is *defining the objective*, exactly what is it that needs to be determined? The need to know whether or not two samples are *different* in character (such as in an ingredient substitution) is a different task from knowing which one is *preferred!* The test objective will determine the type of panelist, methodology, appropriate level of statistical risk to endure and how to interpret results to provide an actionable recommendation. All sensory testing can be defined under two categories: Affective Testing and Analytical Testing.

Definition of "flavor."

Flavor = Aroma + Taste + Texture + Mouth feelings + Appearance + Emotion/Memory

Affective Testing

This includes acceptance/preference testing: Which sample do you prefer? . . . How much do you like it? . . . What don't you like? In order to answer these kinds of questions, true consumers are required. "Consumers" are special individuals who are pre-screened to be actual users of the product tested, they *are not* hungry employees or students who happen to wander in to taste!

Consumer testing is not always the best approach to use. For example, it is an expensive way to determine whether two samples are simply "different", such as in an ingredient reformulation. Nonetheless, some assessment of consumer response to "new" products or "improved" (changed) products is generally advisable.

Affective testing can be performed in-house by employee panelists who frequently are "consumers" of a sort (if they buy and use your food products). They do not necessarily represent your target market, but they are *very useful* for routine directional testing. The key here is: use as many judges as possible ($n° \geq 50$) and repeat the testing, if at all possible. The standard 9-point hedonic scale is very widely used for acceptance testing.

Affective tests include:

Employee Acceptance/Preference—Employees (40-60) are screened for frequent use of the product under study. After a product passes in-house acceptance panels, it may be desirable to submit it to a central location or home placement test to verify in-house results and obtain consumer reaction. A key risk in employee acceptance testing is the bias that employees can have about their products. These biases can be positive or negative and very difficult to estimate and correct. For these reasons, it is recommended that employee acceptance testing be used for screening and general direction only.

Consumer Central Location—These tests are significantly more costly than employee tests, and therefore, should be preceded by adequate in-house testing. Here subjects are screened for their demographic, socioeconomic and product usage profiles. Generally a minimum of 100 consumers

0272-2666/90/0003-3301$04.00/00—

per city, and 2-4 target cities, are recruited for a moderate size test. Since these tests take months to plan, execute and complete, product development direction is slow in comparison to in-house testing.

Consumer tests are recommended for new products, "improved" products, and competitive evaluations. The 9-point hedonic scale, and 5-point diagnostic scales ("too much" . . . "too little") are frequently used.

Focus Group Qualitative Studies seek to develop insight and direction rather than definition or precise measure. Typically, ten respondents who are recruited by screening for specific consumer profiles are led in a 90-120 minute open round-table discussion by a qualified focus group moderator. Feelings and motivations which typically do not surface through quantitative methods (central location tests, home-use tests, mail and telephone surveys) are uncovered.

Focus groups can be "fielded" relatively quickly, their costs are low and they are uniquely flexible as a research technique. They are more difficult to design, execute and analyze than they seem. Their apparent simplicity can lead to sloppy or invalid research. At most, focus groups should only be used to help clarify issues. They should not be used independently or in place of quantitative research.

Home Use Testing requires that pre-screened households actually consume a test product under conditions that approximate normal usage. Respondents can be recruited by telephone, mall intercept, or door-to-door in targeted areas. Products are left with families for a set time period (2 weeks—2 months, depending upon consumption patterns), followed by an interview (telephone, personal or written questionnaire) to quantitate the reactions and opinions of the family to the product(s). In addition to basic preference/acceptance measures, there may be questions about other items such as price, utility and packaging.

Home use testing is absolutely necessary when factors such as ease of preparation, recipe flexibility, potential product abuse, sensory fatigue, etc. are involved. This type of testing is costly and time-consuming (3-6 months to plan and execute).

Mail Panels are actually a lower cost type of home use test for products that are suitable for mail delivery. It permits testing by consumers scattered throughout the country/world under conditions which approximate the norm. Generally, mailing list is developed and indexed by consumer profile (demographics, etc.), then maintained in a database for future testing. Product samples are mailed with the necessary instructions and a questionnaire to be completed and mailed back after consumption.

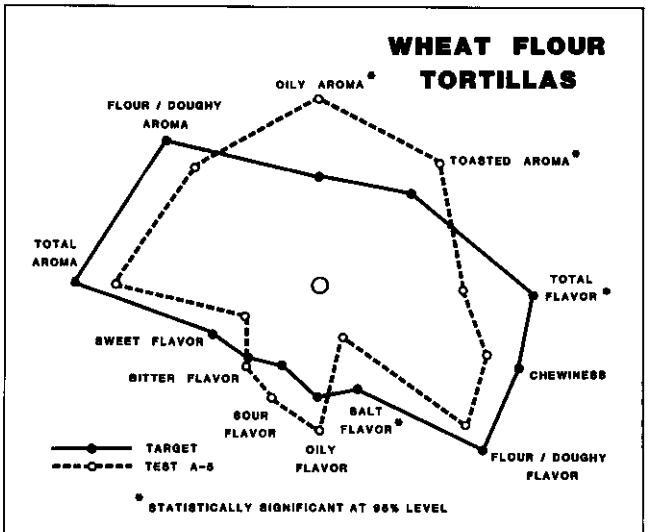

Figure 2: Aroma and Flavor Profiles for Target and Lab Test A-5 Wheat Flour Tortillas.

Telephone Interview. Everyone with a telephone has experienced this popular technique in market research. It is not a good method for obtaining information about a product's sensory characteristics because it is exclusively verbal.

Analytical Testing

This requires the use of objective sensory methodologies using either untrained subjects or trained judges, and helps to *define* the characteristics/properties of foods, but does not define acceptance/preference measures. Analytical testing answers questions such as: Which sample is saltier? Which sample is different? How roasted vs. toasted vs. smokey is the sample? Analytical methods help us define the sensory properties of a food system, but will not directly predict how much it will be liked!

Analytical tests include:

Difference Testing (triangle, duo-trio, paired and multiple comparison, signal detection) is designed to answer the very basic question of: "overall, are the samples different?" Such tests can be conducted with expert and non-expert judges. You can assume that if expert or even experienced judges cannot find a difference between two products, consumers (who are generally less sensitive) would not be able to differentiate.

The triangle test is the most popular of these tests, very sensitive, and easy to administer and interpret. The subject is presented with three samples: two are identical and one is different. Perhaps we are curious about exactly *how* these samples differ in

flavor. Triangle testing will *not* describe the nature of the difference. These tests are designed to simply provide a measure of overall difference, i.e., the samples are different, true or false? If more qualitative information is desired (such as: in which sensory attributes do the products differ?) the Descriptive Analysis method is used.

Descriptive Analysis requires 8-10 expert/trained judges who are thoroughly familiar with the product under study. They know which terms to use for characterizing the product, and how to quantitiate the volume of these sensations within a food. A descriptive panelist does not have supernatural sensory skills, they simply have a focused experience on recognizing and describing flavor attributes. There are a variety of commonly used forms of descriptive analysis, including the Flavor Profile Method (Caul, 1957), Quantitative Descriptive Analysis (Stone, et al., 1980) and the Spectrum Method (Meilgaard, et al., 1987).

Many sensory professionals devise their own customized methods of descriptive analysis yielding the sample type of basic information. If collected appropriately, data from descriptive analysis panels can be analyzed statistically by use of common statistical methods such as the t-test and analysis of variance. Data are often presented graphically in a variety of forms ranging from histograms to circular graphs (Figure 2).

It is important to note that not all differences on the flavor profile are statistically significant and that the relative importance of each attribute may be different. For example, a statistically significant difference between the target and test A-5 (Figure 2) in sweetness would be considered less severe of a problem than a statistically significant difference in bitterness.

Descriptive Analysis is most often used as a technical tool to aid in development or improvement of a product, as well as to delineate problem areas in shelf-life. It is also helpful in understanding the sensory qualities of a product, but it is not the appropriate test to be used when preference or acceptability judgments are required. This technique can be used most satisfactorily, however, in conjunction with hedonic tests to explain affective results.

Ranking is a very straightforward method that can

be used for preference/hedonic measure (i.e., "please rank samples in order of your personal preference") or for analytical assessment (i.e., "rank the samples in increasing order of chewiness"). It is an extension of the paired-comparison (two sample test) approach with multiple samples which the panelist is asked to rank in order according to the degree to which they exhibit some specified characteristic (personal preference or some objective attribute such as "chewiness" or "lemon flavor", etc.).

From 3-10 samples are generally ranked depending upon span of attention/memory as well as physiological considerations. Ease of use is the advantage of this method; its key disadvantage is that rank order results provide information about samples only in relation to each other, with no scaling perspective.

Uses of Sensory Evaluation

An experienced sensory professional can create methodologies and unique, customized approaches to sensory testing. For the beginner, sticking to the basic methods is recommended to insure validity of results. Some of the common applications of sensory testing in the food and beverage industry are presented below:

New Product Development—Unfortunately, most new products are imitations or variations on some established standard. Sensory testing would assure that the standard target has been closely matched or that the unique "point of differentiation" has been made. Descriptive analysis would be the preferred method. Affective testing would determine if the acceptability requirements have also been met. (Note that apples and oranges taste distinctly different, but may have equivalent acceptance scores!)

Product Matching—The objective of product matching is to verify no perceptible difference between the test and standard product. Analytical tests, such as the triangle and duo-trio tests are used.

Product Improvement—Real improvement of a product should be measured by affective tests, to establish whether the experimental product is liked more than the control (i.e., represents an improvement).

Process Change—A process change should maintain or improve the product. The testing sequence is logical:

1. Analytical tests to determine whether the experimental product is different from the control (if it is not different, it cannot be better).
2. Affective tests, if products differ, to establish whether the experimental product is liked as well as or more than the control.

Cost Reduction and/or Selection of a New Source of Supply—A successful cost-reduction program based upon lower-priced ingredients, a lower-cost process, or production in a different location must yield an end product comparable to the product formerly produced. Change to a new supplier of raw materials should also result in an end product comparable to the standard or control. This is a job for analytical tests.

Quality Control—Analytical tests are used during production, distribution, and marketing to ensure that the end product is as good as the standard. Representative samples are usually evaluated as follows:

1. Difference tests to determine whether the sample is different from the standard (if it is not different, it must be as good as the standard.)
2. Descriptive tests, if the sample is found to be different, to indicate how the sample differs from the standard. Results of these tests may be used to guide remedial action, such as changes in processing procedures.

Storage Stability—Analytical tests evaluate product stability during transportation, warehousing and retailing and during storage in the home. To establish information on product shelf life, representative samples are obtained, evaluated initially, and subjected to controlled storage conditions for subsequent tests. At specific time intervals, storage samples are withdrawn and evaluated, generally in comparison with a control. The control must be of the same production lot or batch as the test samples and must be held under conditions known to maintain the original quality. Analytical tests to determine product storage stability may include the following:

1. Difference tests to determine whether the storage samples are different from the control (if no significant difference is found, product stability is assumed).
2. Descriptive tests, used alone or in conjunction with difference tests, to characterize and/or quantify the changes that may have occurred during storage. Descriptive analysis is frequently used in situations where maintenance of a control is unrealistic.
3. Affective tests may be used to determine the relative acceptance of stored products.

Testing Controls

All sensory testing must be conducted under

controlled conditions, following good laboratory procedures. Facilities must be free from mental and sensory distractions, procedures/weights and measure analytically recorded, and the judges should be screened for their basic abilities to smell and taste. Judges should not be burdened with knowledge about the test objectives or samples; subjects must not be biased in any way.

A key consideration in the evaluation of foods is the carrier or media that they are being evaluated with. If the test product can be tasted straight, then it probably should be. If it needs dilution (i.e., a hot sauce, a flavor, a spice or other ingredient), the most precise testing can be done using water as the solvent. Water allows the panelist to most easily detect, separate and quantitate sensory stimuli as no other flavors interfere with, or mask, the product under study. Additionally, any food ingredient should also be evaluated in its final intended application. Food systems will perform and taste differently in fat vs. water soluble basis, in cooked vs. frozen applications, after microcooking vs. frying, etc. Differences in performance vary remarkably with application.

Literature Review

The Institute of Food Technologists has published *Sensory Evaluation Guide for Testing Food and Beverage Products* (1981). This short, precise guide serves as a convenient reference for individuals conducting sensory tests.

For supplementary information on physical conditions of testing, e.g., testing area, sample preparation, and sample presentation, see Amerine, et al. (1965), Eggert (1986), ASTM Committee E-18 (1968; 1973), and Larmond (1977). For information on experimental designs for sensory tests, refer to Cochran and Cox (1957), Amerine, et al. (1965), Winer (1971), and Stone and Sidel (1985). For a glossary of standard definitions of terms relating to sensory evaluation, see ASTM Committee E-18 (1978).

References

Address correspondence to Ms. Marianne Gillette, McCormick & Co. Inc., 202 Wright Ave., Hunt Valley, MD 21031-1066 U.S.A.

MA Amerine, RM Pangborn, and EB Roessler, *Principles of Sensory Evaluation of Food.* New York: Academic Press (1965)

ASTM Committee E-18. *Manual on Sensory Testing Methods.* STP 434. Am Soc for Testing and Materials, Philadelphia, PA (1968)

ASTM Committee E-18. Standard recommended practice for establishing conditions for laboratory sensory evaluation of foods and beverages, E480. In *1975 Book of ASTM Standards,* part 46, p 266. Am Soc for Testing and Materials, Philadelphia, PA (1973)

ASTM Committee E-18. Standard definitions of terms relating to sensory evaluation of materials and products, E253. Am Soc for Testing and Materials, Philadelphia, PA (1978)

ASTM Committee E-18. Guidelines for the selection and training of sensory panel members. STP 758. Am Soc for Testing and Materials, Philadelphia, PA (1981)

JF Caul, The profile method of flavor analysis, *Adv in Food Res,* 7:1 (1957)

WG Cochran and GM Cox, *Experimental Design,* 2nd ed New York: John Wiley and Sons (1957)

J Eggert and K Zook, *Physical Requirement Guidelines for Sensory Evaluation Laboratories,* STP 913. Am Soc for Testing and Materials, Philadelphia, PA (1986)

IFT, *The Wide Scope of Sensory Evaluation,* Inst. of Food Technologists' Sensory Evaluation Div. symposium. *Food Technol* **34**(11): 56 (1980)

IFT, *Sensory Evaluation Guide for Testing Food & Beverage Products,* Sensory Evaluation Division of IFT. pp 50 (Nov 1981)

E Larmond, *Laboratory Methods for Sensory Evaluation of Food,* Pub 1637, rev ed Food Res Inst Canada Dept of Agriculture, Ottawa, Ont (1977)

M Meilgaard, GV Civille, T Carr, *Sensory Evaluation Techniques* Florida: CRC Press (1987)

EE Schaefer, *Manual on Consumer Sensory Evaluation,* STP 682. Am Soc for Testing and Materials, Philadelphia, PA (1979)

H Stone, JL Sidel, *Sensory Evaluation Practices,* New York: Academic Press (1985)

BJ Winer, *Statistical Principles in Experimental Design,* 2nd ed New York: McGraw-Hill (1971)

S Wu, *Product Testing with Consumers for Research Guidance.* STP 1035 Am Soc for Testing and Materials, Philadelphia, PA (1989)