Marketing is concerned with all aspects of consumer acceptance including demographics, psychographics and purchase motivation. Market testing is done at the consumer level in natural settings.

R&D is concerned with developing products having acceptable characteristics at a reasonable cost. Testing usually must be done in "unnatural" settings (i.e., the laboratory). R&D operates on a fraction of the budget allowed marketing and must depend upon marketing for graphics and motivational information. Marketing must depend upon R&D for the development of potentially successful new products.

Product acceptance is the prime concern of both marketing and R&D. Marketing can afford the time and expense of measuring acceptance at the consumer level. R&D has had to develop less costly and quicker methods of evaluating products during the development process. In order to be validly predictive these R&D methods must eventually be correlated with field consumer studies. Considerable care must be exercised in the selection of sensory methods and panels, control of testing conditions and interpretation of data.

It is not difficult to interpret data from field tests as being directly related to consumer acceptance since the test subjects are "real" consumers. It is difficult to relate laboratory-generated data to consumer acceptance since the subjects cannot qualify as representative of the real consumers. This requires a complete understanding of sensory test methods, their application to specific problems, and the meaning of data obtained. It is no wonder that marketing people shy away from the sensory testing area with the explanation that they don't understand the implications. To understand the implications would require considerable more study than most marketing people are willing to expend in this area. R&D on the other hand with its much smaller budget and the problem of time is forced to utilize sensory methods.

IMPORTANCE OF CONSUMER ACCEPTANCE

The basic concern in all sensory testing is product acceptance. Initially we want to know whether or not our product is acceptable to the consumer, or whether the acceptability of our product "matches" or "exceeds" that of another product. Later we will want to know if we can maintain this acceptability level in production through control of ingredients and process (Quality Control). Eventually we will want to know the age at which our product is no longer acceptable to the consumer (Shelf Life).
Acceptance Measurement

The most accurate measurement of consumer acceptance is the sales dollar. Unfortunately this measurement has several serious drawbacks: it is very expensive, it requires considerable time, it is not easily controlled, and it is frequently not reproducible. Measurement of the sales dollar is impractical in most product development situations.

The most practical approach to predicting consumer acceptance today is through the use of sensory panels. Sensory panels measure human responses to sensory stimuli in food products. Consumers perceive product characteristics through the senses of smell, taste, touch, sight and sound. Sensory cues along with behavioral influences provide the consumer with a basis for a judgmental value of acceptance or rejection. Certainly our government and news media are making the consumer very much aware of the sensory characteristics of all products--food and non-food.

Some of the physical characteristics of food products can be measured instrumentally, but we cannot limit our measurement to instruments because it is not possible to specify the interaction of the varying physical properties in producing an overall judgment. Instrumental readings must be correlated with human responses if they are to have any meaning in evaluating sensory characteristics. Instrumental measurements have no real meaning without correlation with human panels. Human measurements can have real meaning without instrumental correlation.

For obvious reasons of better control it is desirable to work towards greater use of instruments. At the present time we must depend upon sensory tests supported by objective measurements wherever possible.

Non-Sensory Factors

Factors of price, availability, utility, convenience and function may be more or less important for a particular product. The importance of sensory characteristics as compared to non-sensory factors should not be underestimated. Consumers expect food products to have acceptable characteristics of taste, smell, feeling and appearance. Ralph Nader recently dismissed flavor as the "deterioration of a food product in the human mouth". Apparently he misses the point that without acceptability on the basis of sensory properties as well as function, the nutritive qualities of a food may remain unutilized. This has been the sad lesson of some attempts to improve the nutrition of the undernourished of the world without regard for their acceptance standards.

Sensory evaluation is the technical approach to the prediction of consumer acceptability. Laboratory tests can be used to determine important product characteristics; to set probable limits of their detectability; and to estimate acceptability. Finally, field tests must be run and the data compared with that from laboratory tests to ascertain the assumed correlation.

SENSORY TEST METHODS

Basically there are three types of sensory methods:
Affective Discriminative Descriptive.

This classification is on the basis of the "task of the subject". Is his task to report his feelings or emotions? to discriminate between products? or to describe the characteristics which he perceives? These categories can be sub-classified by what the subject represents—an instrument or a consumer. If he represents an instrument, the subject is a substitute for a physical measuring technique. If he represents a consumer the subject represents some cross-section of a target population.

Affective Tests:

This category includes both preference and acceptance tests. The terms "preference" and "acceptance" are often used interchangeably. For purposes of product development it might be well to point out that one product may be preferred to another without being an acceptable product.

Affective tests used in product development include ranking, paired comparison, hedonic scales—word and facial, and word action rating scales.

Ranking tests are fast, easily applied and easily interpreted. Their main disadvantages are that they give no indication of magnitude of preference and they evaluate samples only in relationship to each other. Paired Comparison tests can be used with a scale, or in a series in more complex designs. They show a relationship only between samples in a pair or set. Hedonic Rating Scales are very widely used. They indicate magnitude of like-dislike which permits comparison of several samples. Interpretation of terms used and scale values determined may be variable according to product category. Food Action Rating Scales ("I would eat this every opportunity I had— I would eat this if I were forced to") come closer to measuring true acceptance than the Hedonic Rating Scales.

Considerations for using affective tests in product development include the following:

1. Panelists should be persons who normally eat and enjoy the product.

2. Panelists should not have had any prior experience with the product. They should not be persons who have participated in the development of the product.

3. The minimum number of panelists is 24. (Twelve persons making two judgments each do not constitute 24 independent judgments.)

4. Panelists should not be asked to make practical decisions such as that posed by the "Would you buy it?" question. They are not in a practical situation and are not able to give a valid answer to the question.
5. Laboratory preference panels are valid for determination of direction of preference only, not for magnitude of preference. Field tests must be run to make certain of the assumed correlation.

6. "No preference" votes should be considered in the final interpretation of test results and should not be discarded. Discarding of these votes will distort the test findings.

7. Some decision should be made as to the type of judgment required--single or multiple use.

Affective tests determine preference or acceptability. They generate little information concerning flavor construction or flavor difference. Two products may have identical preference ratings and radically different flavor construction. Or a food developer may know very well that a competitor's product is preferred to his own. What he needs to know to improve the acceptability of his product, is "how" the flavor of the two products differs. The only efficient way to obtain such information is to use the appropriate method; in this case, a descriptive type. All sensory testing including affective testing is wasted if the method does not supply the information required.

Affective tests are very important in product development. They are used to determine consumer likes and dislikes at the concept stage and throughout development. They are particularly important at the concept stage when there is no prototype. (If our goal is a product already in existence, we need utilize descriptive techniques only.) If we change our product by using an alternate ingredient or process, we must use affective tests to determine consumer sensitivity (or tolerance) for variations in product characteristics. Shelf life is determined as that age at which our product becomes unacceptable to the consumer.

Discriminative and descriptive measurements must be correlated with affective tests before they can be assumed to have any valid relationship to acceptance. Once this has been accomplished, discriminative and descriptive tests can be efficiently utilized with only occasional checks with affective tests.

**Discriminative Tests:**

Discriminative tests determine difference or similarity between products. They include the standard difference tests (forced-choice) and the sensitivity tests (threshold, dilution).

Basically all discriminative tests measure differences in terms of threshold. If we are using difference panelists as being representative of the consumer, we will select them on the basis of product usage or some other consumer criteria. If we are using difference panelists as physical measuring instruments then we would select them on the basis of "sensitivity to", or "low thresholds for", the flavor difference. In either case the findings must be correlated with data from actual field studies. The dangers of depending upon laboratory difference panels as representative of the consumer population without field correlation are
obvious. It is doubtful that any laboratory panel could reliably reproduce the spectrum of threshold levels existing in the real consumer world. It is generally assumed that laboratory panels are more sensitive than consumer panels because of selection and experience. The reverse may be true depending upon the individual thresholds of the panel members and their representativeness of the consumer population.

The standard difference tests are the triangle, duo-trio and paired comparison. They are precision methods and lend themselves to statistical analysis. They are applicable when differences are slight. When differences are large, the rating scale approach is more useful. They can be used only when intersample effects (contrast, carry-over, etc.) are minimum. (Products with lingering bitterness, smoke, sweetness or tactile factors do not lend themselves to precision discriminative tests.) They should not be used when treatments differ in kind rather than in degree. For example, they should not be used to determine differences between brands of product, since two brands may vary in several dimensions. They can be used to determine difference between two levels of the same ingredient in the same product.

In product development we often have the task of evaluating similar ingredients to perform the same function. For example, we may wish to evaluate several different sweeteners in a product. While the primary function of the ingredient is "to sweeten", any particular sweetening ingredient can affect other flavor characteristics (mouth feeling factors, aftertaste, othertastes, aromatic notes) also. These additional flavor effects can preclude the use of the precision discriminative tests. In such cases we are limited to a single sample test, usually one of the descriptive techniques.

The practice of using scales of "degree of difference" is widespread. The meaning of these scales is questionable. Without anchored scales, i.e., scales with physical standards present, it appears that degree of difference has little meaning. A more meaningful measure would be the percentage of persons finding a difference, and a minimum panel size of sixty persons is required for this measurement.

The design of a triangle difference test can affect the sensitivity of the test. The design allowing for two controls and one experimental sample appears to be more sensitive in some cases than the design which permits one control and two experimental samples. It is important to inspect data for this phenomenon and to make allowances for it in interpretation. Also, if the test is limited to the more sensitive design, and the panelists are aware of this, it may change the probability of the test.

Difference tests tell only whether or not a particular group of persons did or did not find a difference. They do not usually indicate "how" samples are different. They are not necessarily related to consumer sensitivity.

Difference tests are useful in product development particularly in "matching" products. They can be used to determine whether or not a process or ingredient change has produced a measurable difference in the flavor characteristics of a product. They are useful in evaluating the effect of packaging materials, and in maintaining quality control standards.
Threshold tests determine the minimum detectable concentration of a substance (absolute threshold) or the minimum detectable change in concentration (difference threshold). The standard difference tests can be used, or the sequential-up series in which a gradually increasing series of concentrations is utilized. Both sequential-up and sequential-down (staircase approach) series are considered to be part of the basic psychophysical approach to threshold determination. For purposes of product development the sequential-down series is impractical because of physiological complications in going from stronger concentrations to weaker ones. The sequential-up series is recommended for product development. The level duplicated in several runs may be taken as the threshold level. If the standard difference tests are used, the threshold has been defined as being the point in the series where 75% of the judgments are correct (ASTM, Committee E-18). This apparently permits additional confidence in the threshold level over the generally accepted level of 50%.

Threshold tests can be useful in product development in the selection of materials necessary for a particular function within the product but not desirable from a flavor standpoint, e.g., emulsifiers. If a choice is to be made between two or more such materials on the basis of flavor, the material with the highest threshold value (i.e., requiring a higher concentration for detection) would be the recommended material. Threshold tests can be applied in the same manner for the selection of packaging materials (as can lining materials). Threshold-dilution tests can also be used to establish standards for shelf-life. Aged product can be prepared in various concentrations with fresh product in order to select a standard for the shelf-life limit. This approach is particularly useful with beverages or products which blend physically without difficulty.

Threshold tests can also be used for the evaluation of flavoring ingredients in a test medium. Difference threshold tests can be used as quality control checks on different lots of the same flavoring material.

Descriptive Tests:

Descriptive tests answer the question "What are the characteristics of this product?" If the subject is used as an instrument, descriptive methods focus on the product. If the subject is used as a consumer, descriptive methods focus on the consumer's awareness of, or sensitivity to--product characteristics.

Descriptive tests are usually both qualitative and quantitative. They yield a simple description of product characteristics and relative intensity values. Specific techniques for descriptive analysis of flavor and texture have been developed. These are the profile methods. They are based on the use of reference standards and standardized evaluation procedures. Simple rating scales are used for intensity. Flavor profile includes a technique for measuring the duration of flavor notes called "time intensity".

The profile methods utilize small trained panels of four to six persons. They are versatile and can be applied to many different products and problems. They are efficient in that they evaluate many dimensions at the same time.
The profile methods are better known at the research level than in marketing. A descriptive approach more popular with marketing people is the semantic-differential. Semantic-differential is designed to find out which descriptive terms really describe the flavor of a food. A list of terms which might apply to the product is presented to the subject whose task is to select those he feels do apply to the product. In addition he may be asked to assign an intensity value usually on a bi-polar scale. Semantic-differential tests are usually done in the field.

Confidence In A Relationship

Descriptive methods do not lend themselves to ordinary statistical procedures. Data must be accepted "as is" without the security of confidence levels. The importance of descriptive methods has been growing, particularly in the area of subjective-objective correlation. Flavor profile is the only sensory method useful for correlation with instrumental flavor readings (G.C.). In a recent manual on this topic the ASTM refers to a "confidence in a relationship between descriptive data and instrumental data in lieu of a statistical relationship. This approach to interpretation of descriptive data poses no problem for those who rely on a "common sense" approach to the interpretation of sensory data. The profile methods are reproducible and, when validly used and interpreted, can give meaningful data without the use of the usual statistical methods.

In the profile descriptive methods the subjects are used as instruments and are not intended to be representative of the consumer. In the semantic-differential approach the subjects are used as consumers and are considered to be representative of some target population. Problems of dealing with semantics are compounded in the semantic-differential where terms are not usually generated by a panel of persons working with reference standards and standardized procedures. Several persons may use the same term with different meanings (bitter, sour, rancid) or they may use different terms but mean the same thing. The value of using descriptive terms generated by a panel with a common vocabulary is obvious. The best approach in product development would be to use a trained descriptive profile panel to generate terms for later use in semantic-differential approaches at the consumer level.

Advisory Role In Product Development

Another advantage of the profile approach in product development is that the panel is trained to evaluate "blend" and other general characteristics of "good" flavor such as "lack of afterflavor". The profile panels can suggest specific changes in the product for the purpose of improving the flavor. Many flavor profile panels play a very active advisory role in product development.

Descriptive profile methods are primarily used as tools in product development. The panels are small and usually accessible to the product development activity. They are useful in the solution of many product development problems:

1. To match product—particularly when in-depth knowledge of product characteristics is required.
2. To develop new product (no prototype)--useful in initial stages and to assess general characteristics.

3. To change process--descriptive methods can determine if flavor or texture of product is changed, and if so, in which specific characteristics. This approach is especially useful when the size of equipment is changed as in going from lab equipment to pilot equipment, or from pilot equipment to full plant production. Any changes occurring in the product can be described, and the formulation or process changed accordingly to match the product as originally developed and tested at bench top level.

4. To replace ingredients--descriptive data will indicate if original specifications have been maintained, and if not, specifications of changed product.

5. To define specifications of product created--descriptive is the only method which describes flavor as perceived by the consumer.

6. To evaluate effects of packaging materials--descriptive methods give in-depth information about characteristics changed or introduced by packaging materials.

7. To establish and maintain quality control standards--tolerances can be determined by acceptance techniques and maintained by descriptive techniques. Descriptive is the only method which produces a permanent record of the flavor characteristics of a product.

8. To determine shelf-life--descriptive methods can be utilized once acceptance tests have established shelf life limit. The descriptive method is the only one which permits a permanent record of the product at zero time.

Descriptive tests provide information concerning product characteristics. They focus on the product or on consumer awareness of product characteristics. They do not measure preference or acceptance. They are discriminative in that difference can be inferred by comparison of several product profiles. In this case not only do we know that there is a difference but we also know how to describe the difference.

WHICH METHOD SHALL WE USE?

Acceptance is the most important factor in product development but it is often impossible, inefficient or too expensive to use acceptance tests for all product development problems. The use of specific methods for specific P.D. problems has been spelled out.

Just as there is no single sensory method which solves all problems satisfactorily, there is no set order of testing methods. We may
begin with affective and descriptive studies at the concept stage, and use all three types of methods throughout the development program.

It is a fairly common practice to combine affective and difference tests. It is a well established fact that there is a preference bias for the control in this situation. This practice should be avoided.

FACTORS AFFECTING METHODOLOGY

Factors influential in the efficient selection of test methods can and do exist outside the sensory group. Some of these factors are described below.

1. Facilities and Personnel Available.

Methods available will be definitely limited by this factor which is a direct result of management interest and available funds.

Product activity today would be severely hampered by the lack of sensory testing facilities and knowledgeable personnel. Fortunately most food companies do support the sensory testing activity. Some have attractive facilities not fully utilized because of inadequate or untrained staff.

A very big problem for some product developers is the restriction of panel selection and size because of the few persons available. The smaller descriptive panels are often worth the training investment for product developers in this situation. A combination of in-lab descriptive testing and outside consumer studies would provide answers to all questions arising during the development of a product.

2. Time Accounting Procedures.

In many companies it is the policy to charge separate projects for panelist's time on methods requiring longer evaluation sessions as in descriptive tests. Tests requiring less time on the part of the individual panelist (preference tests) are charged to an overhead project. The cost of a descriptive test requiring five panelists for an hour as compared to a preference test requiring thirty panelists for ten minutes each involves the same total manhour expenditure. In the case of the descriptive panel, the product developer's project will be charged; in the case of the preference test, the overhead project will be charged. Administrators will inform you that the company pays for it one way or another. The actual fact is that the effect on methodology can be significant, and can undermine a good all-around sensory testing program.
3. Problem Definition-- Or Lack of Problem Definition.

There is a tendency on the part of product development requesters to order sensory tests as they do groceries--they would like so many preference tests, or so many triangle tests. If the organization maintains any kind of respectable sensory testing group, the requester should define his product development problem to the sensory testing supervisor. The selection of test method should be left to the supervisor.

Many times tests are requested which will not provide the answer to the basic problem. The results are frustrating and demoralizing for all.

Step One for the product developer when considering sensory methods is TO DEFINE THE PROBLEM. The requester must do some thinking about what it is he wants to know before requesting sensory test information.

4. Number Happiness.

Another factor which can affect methodology is "number happiness". There is a tendency to prefer those methods which give us numbers--sometimes without real meaning.

What is the meaning of a 4.5 on a 9-point hedonic scale if a subject will eat a meat product with that rating, but push away a dessert product with the same rating? Or, what is the basis for "killing" a high protein beverage product with a 3.6 hedonic rating while a very popular non-protein beverage enjoyed a 7.6 hedonic rating? The two products are designed for vastly different purposes, but the low hedonic rating of the protein drink caused the death of what might have been a successful product for a specific market. There are probably many financially successful products on the market with low hedonic ratings--but filling the need for certain functions or economic categories.

In our rush to generate numbers we must remember that conditions and need can affect the interpretation of data.

When properly applied, sensory evaluation methods can provide practical information at a reasonable cost. This is very important in product development, particularly at the research level. The relationship of new product development to profit growth and the high mortality rate of new products means that a tremendous amount of product development activity must take place. Sensory evaluation methods utilized in R&D can and do play an important role in increasing the likelihood of product success.
REFERENCES


R. B. SLEETH: Thanks very much Barbara. We hope that for our next event we have psychologically motivated you to triangle test on orange juice, coffee and I hope donuts, that Professor Palmer has in the back for us. So we will see you back at 10:15.

AUTTIS MULLINS: Just two or three announcements. I have been asked to remind you that you are in Florida and that it does require a reconfirmation on your flights. Those of you who are not accustomed to that, when you are leaving Florida you must reconfirm your airline or you might not have your seat. So do that and if you have special transportation problems you can check at the registration desk and have this taken care of. Another announcement -- the Intercollegiate and young Youth Activities Committee members are requested to meet at a table at the noon luncheon today, apparently to transact a little business during the lunch because we don't want you competing with our program after lunch.

Another announcement -- the Meat-judging Team, coaches and other interested people are to meet in Room 361 this evening at 9:00 P.M. or as soon as you return from the picnic. Another announcement regarding the Conference Proceedings -- as you know, we do get a Conference Proceedings printed from the conference program each year. This year is no exception. Everything at the conference is being recorded; therefore we do want to record questions that you may have of the speakers. It will be necessary for you to get to a microphone, if you are close to the microphones in the aisle, please come to those. If not, then just stand where you are, repeat
your name, your affiliation and your question and your question will be repeated at the microphone here. We do want to insure that every question is properly recorded in the conference proceedings.

-- Refreshments --

R. B. SLEETH: In our opening remarks we alluded to providing the consumer with new and improved products, more information relative to its safety and educational material on more desirable methods of preparation. Many of our trade associations are providing the mechanism for the dissemination of knowledge about the food industry and its products to the consumer. One such group that has certainly provided much leadership in this area has been the National Association of Food Chains, headed by its President, Mr. C. G. Adamy. Mr. Adamy will now share with us some of his experiences in educating the consumer by speaking on the general subject of Consumer and Consumerism—Attitudes Toward Specific Products and Ingredients. Mr. Adamy.

# # # # # # # # # #