ORGANOLEPTIC TESTS DEVELOPED FOR MEASURING
THE PALATABILITY OF MEAT

BELLE LOWE
IOWA STATE COLLEGE

No one realizes better than the author, that an adequate presentation of scoring cannot be presented in 20 minutes. For each point that is mentioned, several others are omitted. In addition, the points presented cannot be discussed fully. It is hoped, however, that the material that is presented will promote discussion from this group.

The Necessity for Organoleptic Scores

Objective tests are excellent in fields of research in which performance can be determined by objective tests. There are no objective tests for evaluating certain factors upon which acceptability depends in rating the variations between samples of meat or other foods. In addition, it seems unlikely that usable, valid objective tests for evaluating characteristics such as odor and taste of meat will be developed within a short time. Hence, it seems wise to use all objective tests that give useful information in evaluating scoring data and to spend time in elucidation of the complicating factors in organoleptic scoring. Much thought has been put upon the selection of scoring panels and the development of score cards. Food technologists, however, realize there is still much to do along these lines. In addition tests need to be standardized in order to better compare results from different laboratories.

One of the major difficulties in organoleptic tests is the separation of the palatability factors. Students constantly comment on this point when they start scoring. It takes mental discipline and training to separate and fairly evaluate the flavor of a product if the texture is unacceptable or vice versa. Even one palatability factor, that of taste, may be complicated. There are four taste perceptions, bitter, sour, salt, and sweet. We are further told that all four tastes are intimately connected and that the presence of any one taste affects and may modify the other three tastes (Moncrieff, 1946. p. 12).

Crocker in a study published in 1948 states that raw beef when cold has only a weak odor. The taste is sweetish, salty, and generally bloodlike. When the beef is warmed, considerable odor develops. It is alkaline and smells like a mixture of amines, including piperidine. It is also somewhat eggy, indicating the presence of sulfur compounds. The taste is mild, pleasant, persistent, and chiefly sweet and salty. Crocker found most of the meat flavor came from the meat fiber. With longer cooking the flavor increased, reaching an optimum in 3 to 3.5 hours of boiling, then with longer cooking the flavor decreased. Crocker concluded that cooked-beef flavor is quite complicated chemically, and consists more of odor than taste. Hydrogen sulfide, amines of several kinds including a low simple form, one of the piperidine type, and possible indole were present. The vegetable type odors imply the presence of volatile acids. The only elements of true taste noted were sweetness and saltiness, no doubt owing to blood salts. Astringency was noted in beef extracts as well as in cooked fiber. Many of these constituents are associated in our minds from our chemistry days as unpleasant odors. Yet the small concentrations present and the combination bring about the fragrant, appetizing
odor of cooked meat. None of Crocker's work included flavor developed on the surface of the meat during frying, roasting, etc. Lamb, pork, and chicken all differed from beef in certain flavor characteristics.

Enough of Crocker's results have been included to indicate that scoring meat flavor is not a simple process, and that the flavor is developed to an optimum during cooking, then decreases with longer cooking. Since Crocker (1948) found that the flavor was increased by cooking, it follows that the characteristic flavor of meat is a breakdown product by cooking or is released from meat upon cooking. Work at the university of California (Bouthilet, 1949) indicates the flavoring property in chicken meat is a weak acid produced by cooking.

The Score Card

The score card should be prepared with care and above everything should be simple. A very complicated score card only frustrates the scorer and all too often ends with no significant variations having been found. The number of factors to be scored should be kept as small as possible in relation to the information sought in the study. If four palatability factors are used, they may be aroma, flavor in its wider sense, tenderness, and juiciness. Flavor of fat and texture may be added, although some cuts of certain meats such as veal have little fat for scoring.

If flavor is to be divided into intensity and desirability, then this is difficult scoring and at least two or three palatability factors should be eliminated (such as texture and tenderness or juiciness) to give the scorer time to concentrate on the points being scored.

It is unfortunate from some standpoints that numerical ratings must be assigned to palatability factors for acceptability is not basically a mathematical function. A score of 4 does not imply that the sample is twice as palatable as one with a score of 2, nor that it is two-thirds as acceptable as a sample with a score of 6. But, numerical ratings are practically our only means of evaluating the results statistically and for making comparisons with other laboratories.

Nothing is gained by using a scoring system that has a range of numbers beyond the number of gradations of quality that can be recognized. On the other hand using a range of numbers smaller than the range of qualities that can be recognized is also disadvantageous. Since there is a tendency to use only the numerical ratings between the top and bottom scores a range of 0 to 5 with 1 point variation is less than the gradations of quality factors that can be recognized.

A range of qualities is easier to remember if descriptive terms are assigned to each numerical rating. For example, if scores of 0 to 10 are used, 10 might be assigned the word perfect for a description, 9 excellent, 8 very good, 7 high average, 5 average, 4 low average, 3 fair, 2 poor, 1 very poor, and 0 unacceptable. Varying degrees of inedibility can be assigned negative ratings. Some palatability factors have a limited range of qualities that can be recognized, but tenderness, if based upon the number of chews to masticate a standard size sample to a determined end point, has a wide range of tenderness variations.
The Scoring Panel

Variation in organoleptic scoring is admitted. Even a person who is ordinarily an almost perfect scorer may be thrown off base when he has a head cold. Hence, no hard and fast rule can be made concerning the numbers of scores for a panel. Probably four scorers is the minimum number. However, a still smaller panel composed of members who have high sensitivity and ability to differentiate perceptions may be preferable to a large group with less sensitivity. One of the basic tenets of the statistical treatment of the scoring data is that a larger number of individuals are needed for the panel when the variation in scoring is great. A larger number of scorers will also tend to rule out the experimental error of daily individual variation because of physical or mental conditions.

A scorer should have a high degree of integrity, be able to concentrate and develop an objective and scientific attitude towards scoring meat or other foods, and above all be willing to spend time and effort needed for a fair evaluation of the problem.

If the members of the panel are to be selected on their sensitivity of organoleptic ratings, the consensus at present seems to favor using the food to be scored as a basis of selection instead of dilute salt solutions, sugar solution, etc. In panel tasting it is the ability of the scorer to detect a flavor in a mixture of flavors that is important.

Probably the most common test used at present for the selection of a scoring panel is the triangular one. Prospective scorers are given three samples of meat. The samples are cut from a particular muscle, all the same size. The scorer is told that two samples are duplicates and asked to pick out the duplicates. It is advantageous to ask upon what characteristics, such as aroma, taste, after taste, texture, tenderness, and/or juiciness the decision is made. Some persons will try to make the selection on taste alone instead of using all palatability factors. This lack of observation on the part of the prospective scorer can be called to his attention in the preliminary training. In the triangular test a person has to pick the duplicates 5 times out of 7 for the results to be significant. Persons who cannot detect differences in the samples and who cannot duplicate their judgments in picking duplicate samples should not be placed on the scoring panel.

Members of a panel should be given preliminary training in as wide variation in qualities of meat as possible. This variation should preferably be greater than will be met in the study.

Preparation of and Scoring of the Sample

The sample should be presented to the scorer in an attractive manner. All samples are coded, but coding does not mean that the sample should be placed over a large red or black pencilled number or presented in such a manner that the scorer is adversely affected. The scorer's first reaction is usually to the appearance of the sample.

Each scorer is given a sample, as nearly as possible, from the same anatomical location of each cut. This is important for different parts of the cut may vary in composition of fiber and connective tissue. In addition, a center slice has not been subjected to heat during cooking as long as a slice close to the exterior of the cut, particularly in meat cooked rare or medium done.
The temperature of the meat sample should be about the same for each scoring, or the scorer may be handicapped. Odor is more intense when the meat temperature is high. In addition, the ability to differentiate the four tastes (bitter, sour, salt, and sweet) (Moncrieff, 1946, Chapter 5) is doubled on raising the temperature from 10° to 20°C. Acuteness of taste then is fairly steady from 20°C to 30°C, then falls off as the temperature is raised from 30°C to 40°C. In his summary Moncrieff (p 125) gives the optimum temperature for taste differentiation as 20°C to 40°C.

Slices of meat are cut across the fiber and should be of uniform thickness. Preferably, from the standpoint of chewing the sample to determine its relative tenderness, the meat should be sliced on a slicing machine. This gives sections of the fibers and connective tissue with as nearly uniform length as possible.

Samples should be large enough for the scorer to have more than one taste. Crocker thinks it is preferable to wait long enough between bites for the saliva to clear the mouth rather than eat other food or drink water or beverage between samples. It should also be indicated that if another food or a drink is used between samples, it should also be used before starting to score the first sample.

Moncrieff (1946) states that the sweet taste is most easily perceived at the tip of the tongue, the bitter at the back, the sour at the edge, and the salt at both the tip and edge. Crocker (1945) gives a somewhat different distribution of the taste buds. In some areas two or all of the tastes can be distinguished and there is an area a little less than the size of a dime back of the tongue tip in which no taste is perceived. Other areas in which taste buds are located in addition to the dorsal part of the tongue, are the soft palate, the epiglottis, and the beginning of the esophagus.

It is not necessary for us to decide whether Moncrieff or Crocker is right. It is indicated though that in the preliminary training the prospective scorer should not swallow the sample too hastily, but should develop a definite procedure of rolling the sample over the tongue between chews. In addition, to the taste buds being located in different areas, some taste perceptions are identified more slowly than others. The four tastes in order of ease of detection are bitter, sour, salt, and sweet. It is the combined result of all of these taste and odor perceptions that determine the relative palatability of two or more samples of meat. A good scorer develops a procedure to obtain as much information about each sample as possible.

The scorer should not be interrupted during scoring. Scoring should be done in a place free of odors of other materials and foods.

Each scorer, after the preliminary training, works independently.

There is no definite rule in deciding upon the number of samples that can be scored at one time. The senses of taste and smell fatigue easily, hence the number should not be large. The number of samples that can be scored at one time is also related to the number of palatability factors to be scored. If the score card is complicated two samples are sufficient for scoring. If only part of the palatability factors are to be rated, then four or more samples may be scored at one time.
Subjective and Objective Tests

Subjective tenderness and juiciness of the meat ratings may be compared with objective shear force (pounds or ounces to shear a cylinder of a definite size) and press fluid. The amount of press fluid can be determined by the Carver press of the Minnesota pressometer. We have found that counting the number of chews to masticate a standard size sample to a predetermined end point gives a very easy way of evaluating comparative tenderness in different samples. Do not be surprised if objective tests and subjective tests do not always agree. They may be measuring different things. For example, there are some meats, which with the first bite lose a great deal of juice. After a few chews, the meat may become dry in the mouth. The sample may be rated dry on an organoleptic rating, yet give a high percentage of press fluid.

Evaluation of Tests

Usually the study is not complete until the scoring data have been analyzed by some standard statistical methods to test the significance of the data.

Literature Cited

Bouthilet, R.J., 1949. A note on the flavor constituent of poultry meat, Food Tech. 3: 118


CHAIRMAN BUTLER: Thank you, Miss Lowe.

The discussion will be directed by V. K. Johnson of Illinois.

PROF. JOHNSON: Are there any questions that you would like to direct to Miss Lowe?

PROF. WANDERSTOCK: Miss Lowe, do you think that it is advisable, in setting up a palatability panel for a meat study, to have your committee get together and have a trial run?

MISS LOWE: Oh, yes, that is a preliminary training, particularly if they have not been judging, or if you have been judging for thirty years, like I have, and you had a vacation during the summer, you need to refresh before you go in.

PROF. JOHNSON: Anybody else?

PROF. KLINE: Miss Lowe, would you use the palatability committee after they have eaten, or just before they have eaten, or does that seem to make much difference?
MISS LOWE: Most of ours we use at the time that it is very easy; usually about eleven o'clock, because the meat will come out about that time. I rather gather, if I remember Moncrieff rightly, that that is a pretty good time to score.

In the afternoon, we usually score at about four o'clock. There are some that advocate earlier than that. I remember in our first palatability test, after we started our co-operative meat projects in 1926, that they suggested scoring at 9:30 a.m. but I think it has not been borne out altogether that that is the best time to do it.

PROF. JOHNSON: Any other questions?

If not, I will turn the meeting back to Chairman Butler.

CHAIRMAN BUTLER: Thank you, Mr. Johnson.

Prof. Bratzler has been stalling me off about two days on some questions I have been asking about tenderness of meat, and we will now have a discussion by him on "Determining the Tenderness of Meat by the Use of the Warner-Bratzler Method".

# # #