ARE WE MARRIED TO SENSORY PANELS AND SHEAR TESTS?

by
M. E. Dikeman*

At the 1949 RMC, a Cornell scientist asked a speaker, "Do you think it is advisable to have the taste panel get together for a trial run before beginning a meat palatability study?" The Iowa State speaker answered, "Yes, particularly if you have not been judging meat samples previously."

In general, we would like to think that sophistication of meat sensory evaluation has progressed much since that time. At the 1951 RMC, Professor Deatherage presented results of a "Survey on Organoleptic Testing Methods Used in Meat Research" and one of the main conclusions from the survey was that standardized or uniform methods have not been established so that results between institutions can be readily compared. In the introduction of "Recommended Guidelines of Cookery and Sensory Evaluation of Meat" (1977), it states a serious lack of uniformity among institutions for a) cookery methods, b) sensory panel selection and training, c) sample preparation and d) terminology. So perhaps we have not made as much progress towards sophistication of meat sensory evaluation as we would like to think.

Meat Flavor and Aroma

Tilgner (1974) stated that a bibliography of literature up to 1966 on the sense of taste gives 3000 references. In my review of the recent literature on this topic, I concluded that we are married to taste panels for flavor evaluation of meat. Flavor is perceived, it is integrated by the mind and it cannot be defined by any one or even several chemical methods. When meat products are tasted by consumers or taste panelists, they are perceiving differences in *taste* (sweet, salt, sour or bitter), *aroma* and *mouth feel* (hot or cold, spicy, viscous, etc.). Aftertaste is also very important (Caul, 1967).

Recent research (Mabrouk, 1974) suggests at least 15 classes of organic compounds constituting aqueous beef flavor precursors, and, yet fat is considered responsible for flavor differences between species and for deteriorative changes (oxidative degradation) in meat flavor. Objective methods such as peroxide value, Kreis test, TBA, carbonyls, etc., are limited in merit because values only *indirectly* reflect rancidity (Mabrouk, 1974). Such tests may be useful in determining threshold levels of oxidation, etc. Erickson and Bowers (1974) state only one *direct* method for detecting of lipid deterioration. That is to prepare, package and store the prepared food and evaluate it periodically by sensory means. Therefore, both sensory and chemical methods have their value in measuring "off-flavors" caused by oxidative rancidity.

Pangborn (1967) said "gas chromatography can be used to measure volatiles, not flavor." No instrument or combination of instruments has been developed that reflects the sensory responses to the brain (Reaume, 1975). Objective methods cannot be fully utilized until we understand the biochemistry of meat flavor. Meat flavor cannot yet be measured objectively (Henning, 1974). However, objective methods do provide valuable information and as technology advances, objective methods will likely become more valuable. A good example is a beef flavor study by Persson *et al.* (1973) in which they regressed analytical gas chromatography data against descriptive analyses of aroma. For predicting intensity of the "burnt, smoky" aroma, they used two peak heights to produce a regression equation with a correlation coefficient of .89 (figure 1). They stated that gas chromatography might be very useful for quality control procedures but may be limited to only that.

According to several authorities and proponents of food sensory analyses (Larmond, 1976; Caul, personal communication; Pangborn, personal communication), meat researchers are frequently guilty of being "married to *unsound, misleading* and *invalid* sensory tests"; in other words, guilty of using poor methods. We are often guilty of not distinguishing between "laboratory" and "consumer" taste panels. Laboratory panels should be adequately trained and used as an "analytical tool" or "instrument" to yield sensory discrimination or sensory intensity data. Caul (1967) states that a flavor panel should function about like a good cook. The panelists taste and evaluate and decide if the product is bitter and needs salt or is sour and needs sugar. The flavor profile panel has the ob-

*M. E. DIKEMAN
Dept. of Animal Science and Industry, Kansas State University, Manhattan, KS 66506

jective of answering the question, "How does this taste?" The flavor profile panel was developed in the Food and Flavor Section, Arthur D. Little, Inc., and has been in constant use as an *analytical* flavor method. Caul (1967) further states that a laboratory flavor profile panel should not serve as a consumer panel and make decisions regarding likes or dislikes or preferences (hedonic scales). Yet professional wine tasters not only taste but make judgments and decisions. Seldom is hedonic scaling appropriate for laboratory panels because degrees of liking, preferences, and desirability have reliable meaning only among consumers of the commodity under normal conditions of consumption (Pangborn, 1967).

A flavor profile analysis is a written record of both the aroma and flavor of a product, which are examined and tabulated separately. The tabulation considers five aspects: the individually detectable components; the intensity to which each was perceived; the order in which each was perceived; the whole sensory impression of the aroma and flavor; and for flavor, the aftertaste (Caul, 1967). The individually detectable components of flavor and aroma are recorded in chronological order in descriptive terms defined by reference standards. The intensity to which they are perceived is given in numerical symbols based on a scale ranging from recognition threshold, to slight, to moderate, to strong. Aftertaste also is recorded in descriptive terms and includes not only taste but aroma and feeling sensations. A flavor profile panel can yield information for quality control problems, product development and improvement and for tracing flavor problems engendering consumer complaints.

Laboratory evaluation of liking or preference is not representative of consumer reaction. Certainly consumer-type hedonic ratings cannot be correlated with instrumental data (Pangborn, 1967; Noble, 1975; Reaume, 1975). Intensity ratings, but not "acceptability," may validly be related to physical or chemical measures. Conclusions from the laboratory cannot be extrapolated to the consumers table.

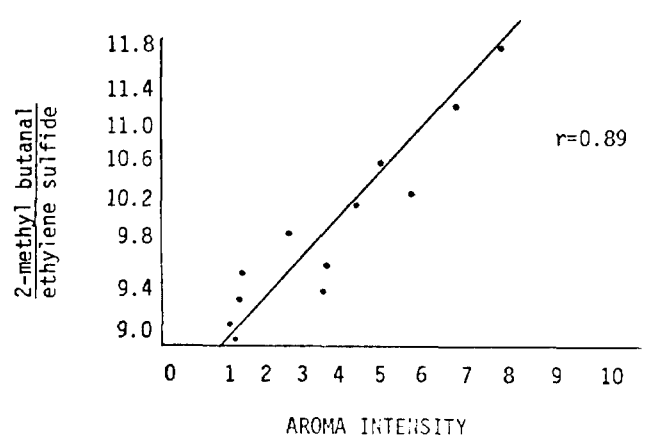

FIGURE 1

Regression line for "burnt, smoky" odor and G. C. data. Persson et al. (1973)

Because sensory evaluation is the ultimate authority in determining food flavor in the first place, and because flavor cannot be evaluated by one or several chemical methods, we *are* married to taste panels for meat flavor and aroma sensory evaluation. We must be sure we are not married to the wrong methods of sensory evaluation. However, we must keep in perspective the tremendous technological advances of recent years and realize that future technological progress may bring us closer to instrumental/chemical methodology for predicting meat sensory characteristics.

Meat Texture

More than any other foodstuff, the literature on meat texture abounds with reports of instrumental/sensory correlations ranging from highly significant to totally nonsignificant. It is important to distinguish between a mechanical test applied to raw meat to predict the tenderness of the cooked product, and an objective method for measuring the tenderness of cooked meat. Despite the progress made so far, much additional work is needed before a satisfactorily predictive test can be established for raw meat (Kapsalis and Szczesniak, 1976). Furthermore, some of the indirect methods of predicting textural properties of meat need extensive work to establish their reliability under all conditions (Voisey, 1976).

In the area of instrumentation, the almost feverish activity of some 20-30 years ago in designing new instruments has now subsided and has given rise to both a critical appraisal of what already exists and a more fundamental approach to understanding what the test measures and what should be measured. That is probably because the more recently devised instruments suffer from much the same problems as the old ones.

Warner-Bratzler Shear Apparatus

Two instrumental methods have predominated in tenderness measurement, the Warner-Bratzler meat shear and the Texture Test System developed by Kramer et al. (1951 a, b). Voisey (1976) discussed criticisms of these two methods. One of the problems with the WB apparatus is that the mechanism moves slowly (22.86 cm min.$^{-1}$) compared with the 150 cm min.$^{-1}$ rates in chewing (Bourne, 1975) so the instru-

ment does not approximate sensory conditions. It would be beneficial if the spring scale deflection could be replaced by electronic force measurement. Standardizing the WB test should be an obvious requirement.

In reporting and interpreting results of WB measurements, test conditions and a typical force-time (deformation) curve should be presented but generally neither is reported. Voisey (1976) also suggested that data cited could include (a) initial yield force representing a tensile rupture; (b) sample area at rupture which indicates compression required to initiate rupture; (c) slope of the force-time (deformation) curve in the second compression phase, which provides an index of firmness; (d) force per unit sample area at the peak; (e) force per unit length of blade cutting the sample at the instant peak force is measured; (f) ratio of the sample area at a selected point (e.g. rupture) to the original and (g) forces and areas in the period following the peak. He further states that standardization of the test blade and slot is lacking, data taken from each sample to generate the tenderness index often are limited, assuming the meat is sheared when it is actually compressed and ruptures under tensile stresses often gives an incorrect interpretation, and consistent terminology that clearly describes the test is absent. Bouton *et al.* (1973) and Cross *et al.* (1973) further stated that WB shear forces correlate poorly with subjective assessment of tenderness when there are large differences in connective tissue strength. Bouton *et al.* (1975a) reported that Instron compression and adhesive measurements reflect changes in tenderness due to connective tissue strength better than shear measurements do.

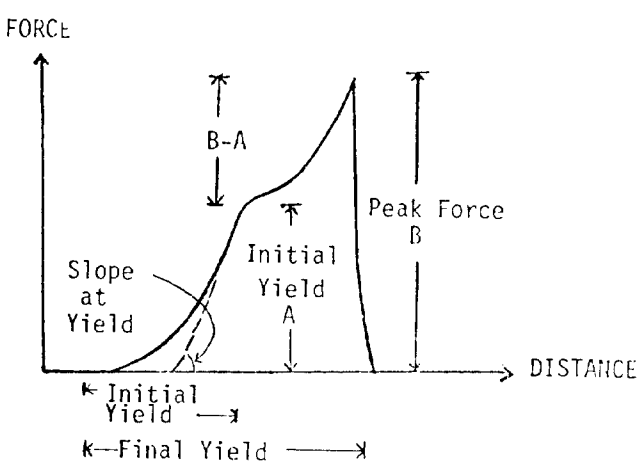

FIGURE 2

A typical shear force-deformation curve for meat (not cold-shortened) showing the various factors measured (Bouton, et al. 1975).

Bouton *et al.* (1975b) studied meat tenderness of samples subjected to various treatments believed to affect different muscle structural components. Their experiments were to determine if differences in the shape of the shear-force deformation curves (figure 2) were well defined, and if so, whether such differences could be used to quantify relative contributions of the various structural components to meat tenderness. Their initial yield force values (for meat that had not cold-shortened) were found to (a) increase with cooking temperatures $\geqq$ 60°C, (b) decrease with post-mortem aging and (c) not significantly affected by animal age. They also showed that the difference between initial yield force and peak force was (a) not significantly affected by post-mortem aging in young or old animals, (b) significantly increased with animal age and (c) significantly decreased by cooking time at 90°C. Their results, plus those of Rhodes *et al.* (1972), suggest that for muscles that have not cold-shortened, a more detailed analysis of the shear force deformation curves could yield some comparative estimates of the myofibrillar and connective tissue contributions towards meat tenderness.

Bouton *et al.* (1975c) reported that nearly 80% of the total variation in tenderness could be accounted for by three measurements: Instron compression for connective tissue toughness, W. B. Shear for myofibrillar toughness and cooking loss for juiciness. Their studies not only dealt with predicting meat tenderness but *why* tenderness can or cannot be predicted. Meat researchers often assume that shear value (actually tensile rupture) alone is a satisfactorily measure and do no further investigating. Certain other chemical or instrumental methods such as myofibrillar fragmentation index (Olson *et al.*, 1976), collagen solubility, etc., help explain *why* there are tenderness differences.

Texture Test System (Kramer Shear Press)

The Texture Test System, which was originally called the Kramer Shear Press, has many of the same problems of other devices which I will not repeat here. The accuracy of recording and the deformation control are not as good as those of other machines. Data reported are generally restricted to maximum force, which is generally assumed to coincide with sample rupture, which is incorrectly assumed to be a shear failure (Voisey, 1976).

Instrumental Textural Profile Analysis

In a recent review, Breene (1975) concluded that textural profile analysis, originally established with the General Foods Texturometer, is now most popularly used with the Instron. Voisey (1976) thorough-

ly discussed the similarities and differences between the G. F. Texturometer and the Instron. Both involve a number of sound fundamental principles applied in an empirical or imitative way. The greatest contribution of instrumental textural profile analysis is to interpretation, which leads to better definitions of textural properties in physical terms. However, the Instron and G. F. Texturometer are quite different in action that affects the interpretation of the readings.

Sensory Evaluation

Sensory evaluation is the ultimate method of measuring meat texture (Larmond, 1976). One factor that makes comparisons of results from one experiment to another, or from one laboratory to another, extremely difficult, is that sensory judgments are relative. Scores assigned to a sample are influenced by the range in tenderness of samples being judged concurrently. Moskowitz and Kapsalis (1975) suggested that simulated foods be developed for use as standards for texture. Stanley *et al.* (1972) provided judges with rehydrated, unflavored, texturized soy protein as reference samples and instructed them to evaluate meat samples relative to the reference sample and not to each other.

Harries *et al.* (1972) examined the interrelationships between the individual parts of a multi-component scoring system. Judges identified seven characteristics as being important in meat texture: resistance to initial chewing, wetness, juiciness, cohesiveness, hardness, overall texture and chew count. Factor analysis indicated that the seven characteristics were not independent but could be reduced to two characteristics: toughness-tenderness and juiciness. Larmond (1976) stated that chew count is not recommended as a standard method.

Cross *et al.* (1977) developed step-by-step procedures for interviewing, screening, training and testing a meat descriptive panel. Factors evaluated were tenderness, juiciness and amount of detectable connective tissue. That procedure requires considerably more time, and therefore expense, than many meat researchers have invested in taste panel procedures.

The first step is the personal interview to determine interest and availability of the potential panelist and what is expected on a taste panel. Screening is the second step and utilizes triangle tests. One of the following decisions is made after each triangle test: accept the candidate as a potential panelist, reject him or continue testing. Training is the third step and is best accomplished through sessions in which various samples of the product types in the tests are evaluated and discussed. For example, steaks from animals of varying ages are used to demonstrate connective tissue differences. Steaks cooked to different degrees of doneness provide ranges in juiciness. In reality, training is never completed.

Testing is the last step. Nine samples are selected to cover the full range of the attribute being trained for. The design is outlined in table 1. Data analysis treats the data for each candidate as a one-way analysis of variance with nine treatments and four observations per cell. With this design, day and session effects can be studied. From the ANOVA table, the F-ratio is calculated as mean square treatment ÷ mean square error. The F-ratio is a measure of a panelist's ability to give different scores to different samples while being able to repeat himself on the same sample later. The degree to which a person discriminates between samples and is consistent in replicate judgments will be reflected in his F-ratio (table 2). Candidates can be ranked on the basis of these F-ratios (table 3).

To determine the preciseness in which a meat descriptive panel can be trained, three additional panels were trained at three different universities using the same techniques described. Procedures and equipment were standardized as much as possible between panels. Common samples were used to test panels during training. Correlation co-efficients were calculated between all panels (table 4). High correlations for tenderness and connective tissue suggest that meat descriptive panels trained by a valid and standardized procedure can yield precise and reliable sensory data. The low variability in juiciness scores resulted in low correlations among panels for that sensory characteristic.

DAY 1			DAY 2			DAY 3		
Session			Session			Session		
1	2	3	1	2	3	1	2	3
S_1[b]	S_4	S_7	S_1	S_4	S_7	S_1	S_4	S_7
S_2	S_5	S_8	S_2	S_5	S_8	S_2	S_5	S_8
S_3	S_6	S_9	S_3	S_6	S_9	S_3	S_6	S_9

TABLE 1

DESIGN LAYOUT FOR A TASTE PANEL TEST[a]

[a]From Cross *et al.* (1977).
[b]S_1 = sample number one, S_2 = sample two, etc.

TABLE 2

SAMPLE TASTE PANEL TEST; F-RATIOS BY PALATABILITY ATTRIBUTE[a]

		F-Ratios										
		Panelist Number										
Attribute	Panel Leader	1	2	3	4	5	6	7	8	9	10	11
Tenderness	14.40	2.76	7.12	7.54	8.29	8.87	8.67	6.12	5.37	6.77	8.78	3.64
Juiciness	5.33	2.37	3.59	8.03	3.58	8.78	6.94	6.14	4.59	1.89	7.50	3.18
Connective tissue	8.73	4.02	4.39	4.07	5.37	5.32	8.32	3.76	0.95	6.84	12.32	4.03

[a]From Cross **et al.** (1977).

TABLE 3

SAMPLE TASTE PANEL TEST; RANKS OF PANELISTS BASED ON F-RATIOS[a]

		Ranking										
		Panelist Number										
Attribute	Panel Leader	1	2	3	4	5	6	7	8	9	10	11
Tenderness	1	12	7	6	5	2	4	9	10	8	3	11
Juiciness	6	11	8	2	9	1	4	5	7	12	3	10
Connective tissue	2	10	7	8	5	6	3	11	12	4	1	9
Sum	9	33	21	16	19	9	11	25	29	24	7	30
Overall rank	2	12	6	5	6	2	4	9	10	8	1	11

[a]From Cross **et al.** (1977).

Juiciness

Taste panel assessment of meat juiciness involves two main effects, moisture released by the meat during chewing and moisture from salivation. The latter is highly subjective and almost impossible to measure objectively; the former is highly related to the moisture content of cooked meat. The high speed centrifugation method of Akroyd (reported by Bouton *et al.*, 1971) to determine the expressible juice in cooked meat (by a water bath method) and measurement of cooking losses (Bouton *et al.*, 1971 and 1974) both correlated highly with sensory panel juiciness scores ($r = .97$ and .92, respectively). Cooking losses from meat conventionally cooked in ovens would probably vary more.

Meat juiciness can be adequately evaluated by cooking loss, moisture determination in cooked meat or sensory panel methods. It seems that juiciness is the easiest of meat sensory attributes to evaluate. Training a sensory panel to evaluate juiciness would be much easier than training a panel to evaluate flavor, and cooking loss or moisture content is relatively easy to measure.

Summary

Meat flavor cannot presently be evaluated by chemical or instrumental methods. Certain chemical and instrumental methods, such as gas chromatography, are useful in quality control to monitor threshold levels of certain volatile constituents. Progress in technology will likely result in chemical/instrumental methodology that will be more useful in predicting meat flavor and aroma. Because sensory evaluation is the ultimate authority in determining meat flavor, it is apparent that we must rely on taste panels for meat sensory evaluation. However, laboratory panels should be adequately trained and used as an "analytical tool" or "instrument" to yield sensory discrimination or sensory intensity data.

For meat texture evaluation, we are not married to either sensory evaluation or instrumental methodology. Properly trained and "calibrated" taste panels

TABLE 4

CORRELATION COEFFICIENTS AMONG TASTE PANELS AT DIFFERENT UNIVERSITIES FOR THREE PALATABILITY ATTRIBUTES[a]

Panel Number	TENDERNESS			
	Panel 1	Panel 2	Panel 3	Panel 4
Panel 1	1.00	.88	.93	.92
Panel 2		1.00	.90	.90
Panel 3			1.00	.94
Panel 4				1.00
	JUICINESS			
	Panel 1	Panel 2	Panel 3	Panel 4
Panel 1	1.00	.17	.17	.42
Panel 2		1.00	.07	.30
Panel 3			1.00	.36
Panel 4				1.00
	CONNECTIVE TISSUE			
	Panel 1	Panel 2	Panel 3	Panel 4
Panel 1	1.00	.86	.81	.82
Panel 2		1.00	.84	.85
Panel 3			1.00	.79
Panel 4				1.00

n = 55 **longissimus** steaks.
[a]From Cross **et al.** (1977).

yield analytical data and represent the ultimate authority for meat texture evaluation. However, several instrumental methods can yield analytical data quite adequate for the objectives of many research studies. These instruments must be calibrated, used correctly and the data must be accurately and completely interpreted. Future technological advances hold considerable promise for new approaches to instrumental measurement of meat texture and likely will be the predominant method in the future.

BIBLIOGRAPHY

AMSA Meat Cookery and Sensory Guidelines Committee. 1977. Recommended guidelines for cookery and sensory evaluation of meat. American Meat Science Association, Chicago, Illinois.

Breene, W. M. 1975. Application of texture profile analysis to instrumental texture evaluation. J. Texture Studies. 6:53.

Bourne, M. C. 1975. Interpretation of force curves from instrumental texture measurements. Rheology and Texture in Food Quality. AVI Publishing Co., Inc., Westport, Conn.

Bouton, P. E., P. V. Harris and W. R. Shorthose. 1971. Effect of ultimate pH upon the water holding capacity and tenderness of mutton. J. Food Sci. 36: 435.

Bouton, P. E., A. L. Fisher, P. V. Harris and R. I. Baxter. 1973. A comparison of the effects of some post-slaughter treatments on the tenderness of beef. J. Food Technol. 8:39.

Bouton, P. E., P. V. Harris, W. R. Shorthose and D. Ratcliff. 1974. Changes in the mechanical properties of veal muscles produced by myofibrillar contraction state, cooking temperature and cooking time. J. Food Sci. 39:869.

Bouton, P. E., P. V. Harris and W. R. Shorthose. 1975a. Possible relationships between shear, tensile and adhesion properties of meat and meat structure. J. Texture Studies. 6:297.

Bouton, P. E., P. V. Harris and W. R. Shorthose. 1975b. Changes in shear parameters of meat associated with structural changes produced by aging, cooking and myofibrillar contraction. J. Food Sci. 40:1122.

Bouton, P. E., A. L. Ford. P. V. Harris and D. Ratcliff. 1975. Objective-subjective assessment of meat tenderness. J. Texture Studies. 6:315.

Caul, J. F. 1967. The nature of flavor. Cereal Science Today. 12(7):273.

Cross, H. R., Z. L. Carpenter and G. C. Smith. 1973. Effects of intramuscular collagen and elastin on bovine muscle tenderness. J. Food Sci. 38:998.

Cross, H. R., R. Moen and M. S. Stanfield. 1977. Guidelines for training and testing a meat descriptive panel. J. Food Sci. (accepted for publication).

Deatherage, F. E. 1949. A survey of the organoleptic

testing methods used in meats research. Proc. Recip. Meat Conf. 3:184.
Erickson, D. R. and R. H. Bowers. 1974. Objective determination of fat stability in prepared foods. Objective Methods for Food Evaluation; Proc. of a Symposium, 133.
Harries, J. M., D. N. Rhodes and B. B. Chrystall. 1972. Meat texture. I: Subjective assessment of the texture of cooked beef. J. Texture Studies. 3:101.
Henning, G. J. 1974. Evaluation of fat stability in emulsified foods. Objective Methods for Food Evaluation; Proc. of a Symposium, 155.
Kapsalis, J. G. and A. S. Szczesniak. 1976. Instrumental testing of meat texture—comments on the past, present and future. J. Texture Studies. 7:109-113.
Kramer, A., G. J. Burkhardt and H. P. Rogers. 1951a. The shear-press, a device for measuring food quality. Canner. 112:34.
Kramer, A., K. Aamlid, R. B. Guyer and H. Rogers. 1951b. New shear press predicts quality of canned limas. Food Eng. 23:112.
Larmond, E. 1976. Texture measurement in meat by sensory evaluation. J. Texture Studies. 7:87-93.
Mabrouk, A. F. 1974. Beef flavor precursors and quality assessment. Objective Methods for Food Evaluation; Proc. of a Symposium, 249.
Moskowitz, H. R. and J. G. Kapsalis. 1975. The texture profile: its foundation and outlook. J. Texture Studies. 6:157.
Noble, A. C. 1975. Instrumental analysis of the sensory properties of food. Food Tech. 29(12):56.
Olson, D. G., F. C. Parrish, Jr. and J. H. Stromer. 1976. Myofibril fragmentation and shear resistance of three bovine muscles during post-mortem storage. J. Food Sci. 41:1036.
Pangborn, R. M. 1967. Use and misuse of sensory methodology. Food Quality Control. 15:7.
Perrson, T., E. von Sydow and C. Akesson. 1973. The aroma of canned beef: models for correlation of instrumental and sensory data. J. Food Sci. 38:682.
Reaume, J. B. 1975. Methodology for sensory testing with emphasis on limitations of tests. Proc. Recip. Meat Conf. 28:375.
Rhodes, D. N., R. C. D. Jones, B. B. Chrystall and J. M. Harries. 1972. Meat texture. II. The relationship between subjective assessments and a compressive test on roast beef. J. Texture Studies. 3:298.
Stanley, D. W., L. M. McKnight, W. G. S. Hines, W. R. Usborne and J. M. deMan. 1972. Predicting meat tenderness from muscle tensile properties. J. Texture Studies. 3:51.
Tilgner, I. D. J. 1974. The sensory evaluation of quality. Die Fleischwirtschaft. 8:1332.
Voisey, P. W. 1976. Engineering assessment and critique of instruments used for meat tenderness evaluation. J. Texture Studies. 7:11-48.