Procedures for the Sensory Evaluation of Horticultural Crops[1,2]

Christi M. Heintz and Adel A. Kader
Department of Pomology, University of California, Davis, CA, 95616

While some of the quality attributes, such as color and texture, of fresh horticultural crops can be measured by objective methods, accurate determination of flavor quality requires the use of subjective methods, i.e., sensory evaluation procedures. These include use of laboratory panels for detecting and describing differences among samples and use of consumer panels for indicating quality preferences. The proper use of such techniques involves following specific procedures for each sensory evaluation test as described in this article.

C. M. Heintz

A. A. Kader

Sensory evaluation involves the measurement, quantification, and interpretation of the sensory characteristics of foods and consumer products through the use of human subjects acting as judges. The scope of sensory evaluation lies far beyond fruits, vegetables, meats, cereals, and other food products to include the evaluation of any product a consumer may purchase—ornamentals, mouthwash, floor polish, even deodorants and air fresheners. In the evaluation of food products, the sensory scientist is mainly concerned with an evaluation of the appearance, flavor (taste and aroma), and texture parameters of a particular commodity. Since the publication of the book by Amerine et al. (5), several useful articles and booklets dealing with this topic have appeared, including those published by the American Society for Testing and Materials (1, 2, 3, 4) and the Institute of Food Technologists (8, 9, 10). Stevens and Albright (17) briefly discussed the application of sensory evaluation procedures to horticultural commodities. The objective of this article is to elaborate on proper use, interpretation, and presentation of results of sensory evaluation methods for horticultural crops.

Why a sensory panel?

The results of sensory evaluation usually are but one aspect of a larger array of tests, including chemical and objective tests, which together produce a total evaluation of a specific commodity. While chemical and objective tests provide specific information about the composition and physical nature of the commodity, sensory evaluation provides a means to study integrated parameters and the impact of different variables on the sensory quality of the product. Results of the different sensory, chemical, and objective tests can be easily correlated to identify the relationships among the chemical and physical properties of a commodity and its sensory qualities. For horticultural crops, testing variables may include, for example, a comparison of cultivars, production areas, cultural practices, harvesting procedures, maturity stages at harvest, or postharvest handling methods.

Laboratory sensory tests—information they cannot provide

Sensory evaluation panels can generally be classified into 2 groups: 1) consumer panels and 2) laboratory panels. Consumer panels involve thousands of individuals who are asked for their preference of one product over another. Consumer panelists are chosen from a representative portion of the consuming public for which a product is targeted. Panelists may be selected for their age, sex, socioeconomic level, race, education, or location of residence. The panel must be large enough to overcome some of the extreme variability which occurs among individuals when asked for their preference. Consumer panels are mainly set up by professional marketing research organizations and products tested are the result of months of screening products and treatment variables, then reformulating

[1]Received for publication Nov. 5, 1982.
The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked *advertisement* solely to indicate this fact.
[2]Photographs for Fig. 1–4 were taken by Don Edwards.

Downloaded from https://prime-pdf-watermark.prime-prod.pubfactory.com/ at 2025-07-05 via Open Access. This is an open access article distributed under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/). https://creativecommons.org/licenses/by-nc-nd/4.0/

and retesting in the laboratory. Even these tests may not adequately provide information about products a consumer will actually purchase—final purchasing depends upon a variety of factors, including previous experience with a product, advertising, availability, price, packaging, convenience, and nutritional factors.

In contrast, laboratory panels consist of relatively few (12 to 25) members selected and trained for their ability to discriminate between small variations in sensory characteristics. They cannot be familiar with the experimental variables being tested, yet, they are usually people who work near enough to the laboratory to be available on a daily basis for testing. These individuals are therefore not representative of the consuming public. Further, their screening, training, and experience in making judgments in a typical sensory laboratory—including individual booths, lighting variations, prepared samples, and scoring sheets provide conditions unlike those found in a typical marketplace.

The most common mistake made in a sensory test is to ask a laboratory panel for their preference. Hedonic evaluations (preference tests) are made only after exhaustive laboratory panels, final product screening, and with the use of a very large number of "typical" consumers.

Sensory evaluation procedures

Methods for obtaining accurate and valid data from human subjects include very specific techniques and are to be taken as seriously as, for example, procedures for chemical analyses. Sensory studies must be planned and standardized to be as objective and scientific as possible. Many factors contribute to the standardization of sensory tests: 1) screening and training of panel members; 2) a suitable test environment; 3) techniques for sample preparation and sample presentation; 4) statistically valid methods for obtaining and analyzing results; and 5) standard reporting procedures.

Fresh horticultural commodities present a greater challenge in terms of sensory evaluation procedures than processed products because of: 1) greater inherent variability among individual units of a given commodity; 2) difficulty in making suitable reference samples available to the panel; and 3) difficulty in simultaneously comparing samples picked at different stages of maturity or subjected to various storage treatments. For example, evaluating the effects of different storage conditions on flavor quality of a commodity poses a particular problem. Ideally, tests involving different storage periods should be initiated on different dates so that they terminate the same day (12). For lengthy storage periods, however, sensory evaluation must be carried out on each treatment separately (19). When possible, references may be used to ensure consistency in judging. One approach for providing reference samples is to add specific quantities of sugars, acids, or other chemicals to cut portions of the commodity (11, 18). However, the success of this method is dependent upon the extent and uniformity of incorporating the added chemical into the tissue.

Panel members

Recruiting at least twice as many judges as necessary will account for those individuals who find they cannot spare the time to volunteer. A minimum of 12 to 15 final judges is required. Only those people who seem keenly interested in participating and who will be available for the entire experimental period should be recruited.

On the first day of the panel, the experimenter should spend a little time with each judge, acquainting him or her with the score sheet and scoring system, definitions of any descriptors, procedures for sampling the commodity, and any other special instructions. These instructions and a statement of test objectives could also be accomplished in a group orientation meeting.

Two days to 1 week or more should be spent on training judges, depending on the complexity of the test, followed by 2 to 3 days of preliminary testing using the final scoresheet and number of samples to be presented during actual testing. Training and preliminary testing not only familiarize judges with the sample, descriptors, and scoring system, but also familiarize judges with the entire testing environment and help them to set a daily pattern of participating in the sensory tests.

Training judges should begin by presenting just a few samples, then daily increasing the number of samples to that amount which will be presented during actual testing. If a descriptive scoresheet is being used, begin testing judges on just 2 or 3 of the easiest characteristics, gradually adding more terms and more samples as the days progress. Because judges will be more confident in themselves if they feel they can discriminate samples fairly easily, ensure that samples are distinctively different the first few days of training. For example, if the test involves rating sweetness of canned peaches, allow 1 sample to sit overnight in a canning syrup, altered with the addition of sucrose, so that the sample will taste much sweeter the day of the panel. In fresh fruits, variations in degree of sweetness can be accomplished by presenting fruits of different maturity stages. Providing samples of obvious variation can also help screen nondiscriminating judges. Final panel selection depends on the accuracy and consistency of each panel member. Judges should give consistant results over several replications of one treatment.

In addition to reproducibility of their results, the final panel of judges should consist of members who have shown continued interest throughout the training and screening period and have no problem participating on a daily basis. Judges will more easily remember their commitment to the panel if they are required to come every day rather than 2 or 3 days a week. Time of day probably has little influence on sensory perception; however, morning panels are recommended, as judges seem more motivated to participate in morning than in afternoon panels.

Sensory tests differ from most other scientific tests in their dependence on continued volunteer support. Judges need encouragement throughout the course of the panel. Methods to enhance panel motivation include: 1) short discussions with individual panel members on the particular commodity being tested, i.e., discussing interesting facts or cultural practices about the fruit or vegetable without giving clues to any of the experimental variables being tested; 2) posting interesting news stories, clippings, comics, etc., on a bulletin board outside the booth area; 3) giving judges feedback as to their performance by posting some results; and 4) providing daily treats (cookies, fruits, juice, etc.) in appreciation for their participation.

The sensory experimenter must deal courteously with the panelists. Giving plenty of notice for cancelled or added testing days, posting a calendar of days and hours the panel will meet and adhering to those dates, and excusing judges when they feel ill or have a cold, all show some consideration for the judges. It is helpful to give each judge a small card to post near his or her desk or work area as a reminder of the daily panel.

Test environment

The location of the sensory testing area must be convenient for panel members. It should be clean, comfortable, temperature-controlled, free from distracting visual stimuli, odors, and noises, and generally conducive to concentration. A spatial arrangement designed to promote working efficiency is the main consideration; a large sample preparation area is required with ready access via small sliding doors for passing samples into partitioned booths.

The booth area should be separate from the working area so that judges are not distracted by the noises and odors of sample preparation, and do not observe any clues as to the samples and treatments being presented. Booths must be partitioned from one another to eliminate interaction between judges and to facilitate concentration. Individual booths should provide: 1) lighting which can vary according to the test objective (white incandescent bulbs of standardized wattage for appearance/color tests and red lighting which masks appearance variations for taste evaluations); 2) facilities for oral water rinses and sample expectoration; and 3) a signaling method (i.e., a switch which lights a bulb in the preparation area) for the judge to alert the experimenter of problems, questions, or completion of a test.

Sample preparation

Samples chosen must be free from cracks, bruises, insect damage, and other defects. To reduce variability, samples of a given treatment should originate from the sample production area with similar cultural practices and harvest maturities. Timeliness in handling is required to ensure minimal physiological changes in the commodity.

Prepare samples so that they are of a typical physical state and temperature of consumption when evaluated by the sensory panel.

Downloaded from https://prime-pdf-watermark.prime-prod.pubfactory.com/ at 2025-07-05 via Open Access. This is an open access article distributed under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/). https://creativecommons.org/licenses/by-nc-nd/4.0/

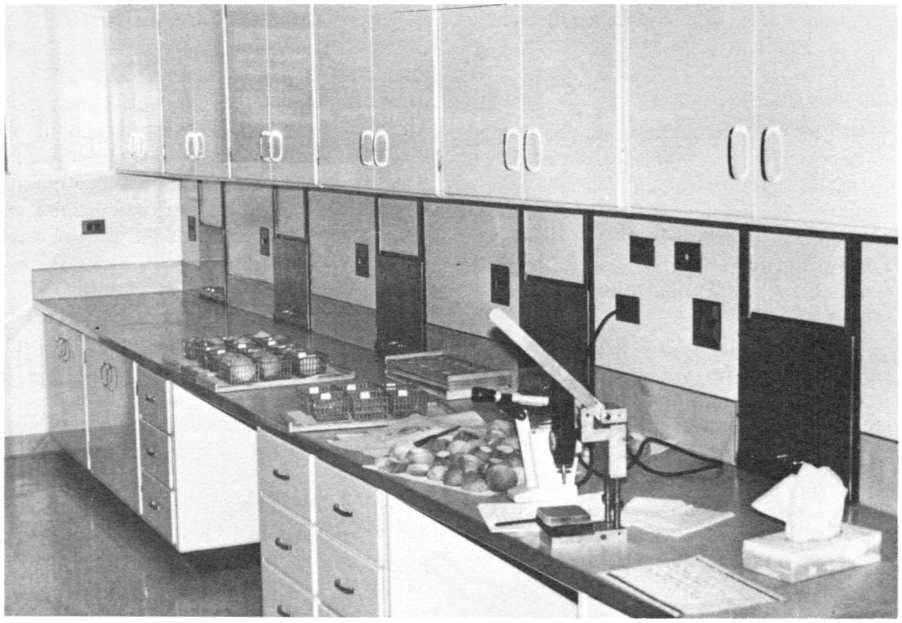

Fig. 1. Sample preparation area; also shown are a UC Fruit Firmness Tester and a refractometer used for determining fruit flesh firmness and soluble solids content, respectively, of individual fruits to be evaluated by the panel.

Allow enough time for cleaning, shelling, peeling, moisture equilibration, slicing, cubing, or pureeing. In addition, allow time for preparation or alteration of samples used in training judges. Procedures such as moisture equilibration, cleaning, and shelling should be completed well in advance of the panel, whereas slicing, cubing and pureeing should be done just prior to evaluation to minimize loss of volatiles, browning or other discolorations, drying out, etc. Deciding the physical state (sliced, diced, pureed, etc.) to present a commodity depends upon the type of commodity, quantity available, typical state of consumption, treatment variables, etc. In sliced, diced, or pureed fruit, the experimenter can discard stem- and blossom-end sections to eliminate large proportions of peel, and combine several fruits per sample to reduce variability due to morphological region of the fruit (e.g., gradations in sweetness from stem- to blossom-end) and variability among different fruits. Diced or pureed commodities, however, are unsuitable for texture and firmness evaluations.

Representative samples should be prepared similarly for corresponding chemical and physical analyses. These analyses should be performed on the same day as the sensory panel so as to minimize effects due to time.

Sample presentation

Judges, being human, use all available information in making decisions, therefore, sample should be presented so as not to provide any information judges may interpret as clues to the identity or nature of a sample. Standardization procedures include: 1) using similar sample containers; 2) providing similar sample sizes; 3) labeling samples with 3-digit random numbers; and 4) presenting samples in a randomized order. Sampling containers can be any material which is easy to use, lable, and clean, and will not impart odors or flavors to the sample. Lids may be necessary if odors from one sample could interfere with the evaluation of another sample. Adequate sample size is necessary for good representation of the commodity.

Judges should be provided with spoons, forks, or knives, water glasses for rinsing, spittoons for expectoration of samples, napkins or tissues, pencils, a set of instructions and definitions, and a proper scoresheet. Instructions will vary with the test but should include reminders to: 1) read the definitions provided; 2) rinse between samples; 3) expectorate samples to avoid fatigue of the senses and any digestion or indigestion variables; 4) sample the reference initially, if one is provided; 5) sample the commodity consistently; and 6) signal the experimenter upon completion of the test. The experimenter wants to ensure all judges take samples in a like manner. For instance, if judges are required to sample a slice from a whole fruit, a diagram should be provided with instructions to hold the fruit with the stem end up, the suture toward the judge, and to cut a "bite-size" wedge from the right cheek of the fruit. If diced fruit is presented in its own juice, the judge should be instructed to sample a representative portion of the fruit and liquid. Specific instructions as to whether the test involves difference or descriptive testing and those scoring methods may be provided on the scoresheet. Procedures followed in the UC Sensory Evaluation Laboratory are shown in Fig. 1–4.

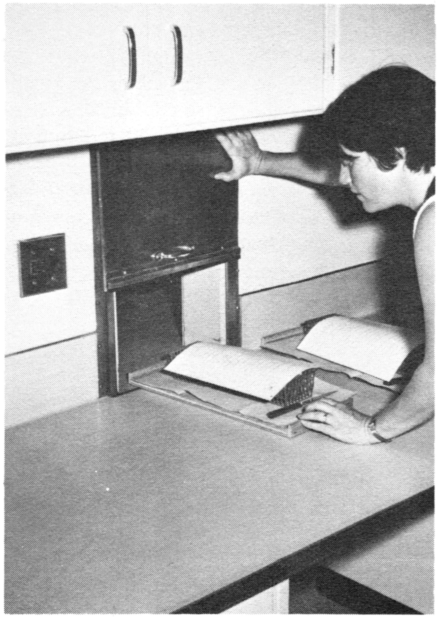

Fig. 2. Sliding door for passing samples from preparation area into the testing booth.

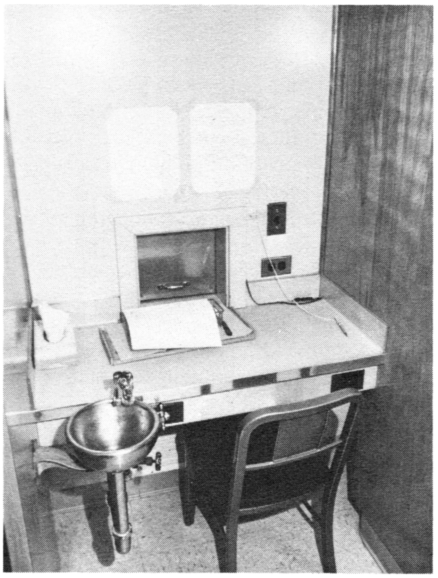

Fig. 3. A sensory evaluation booth equipped with proper lighting, facilities for oral water rinses, and a switch for signaling the experimenter when needed.

Fig. 4. A taste-panel member evaluating peach fruit samples.

Downloaded from https://prime-pdf-watermark.prime-prod.pubfactory.com/ at 2025-07-05 via Open Access. This is an open access article distributed under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/). https://creativecommons.org/licenses/by-nc-nd/4.0/

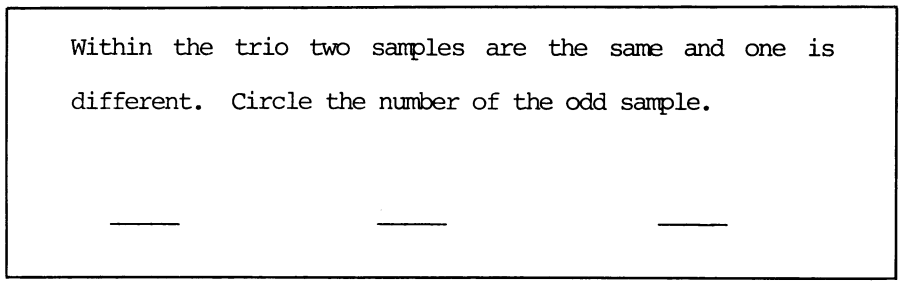

Fig. 5. Example of a triangle test score card.

Difference vs. descriptive tests

Panel members may be asked whether samples simply differ from one another (difference tests) or they may be asked to rate several sensory characteristics of one or more samples (descriptive tests). Triangle tests are one of the most common methods of difference testing and involve the presentation of 3 samples. The judge is then told that 2 of the samples are the same and 1 is different and to choose the odd sample (Fig. 5). In descriptive tests, the judge is given several sensory characteristcs (i.e., firmness, sweetness, sourness, etc.), and asked to rate a sample for each of the characteristics using a numbering or labeling system or marking a line. Our experience has shown us that marking a point on a line anchored near the ends of the line by "less" or "more" intensity of the characteristic (Fig. 6) introduces less bias than using numbers or word lables to which numbers have been assigned. When a point is marked on a line, as in the example, the experimenter then derives numbers by measuring the line in centimeters. While difference testing results in identification only of samples which differ, without providing any information on the sensory characteristics which make them differ, descriptive tests evaluate specific color, texture, aroma, or taste characteristics and provide much more information on the commodity and variables being tested. Several publications (1, 5, 6, 10, 13, 16) further describe the types and usage of difference and descriptive sensory tests.

If descriptive tests are to be used, the proper sensory characteristics must be chosen to derive the most information about the treatment variables. Procedures for initiating a descriptive-type test include preparing samples representative of the different treatments and enlisting the help of 3 to 4 persons experienced in sensory panels and who are familiar with the commodity to choose several sensory characteristics important in describing the treatment variables. It is then helpful to bring the sensory panel together for an orientation meeting and round-table discussion of sensory descriptors. The panel meeting can serve to choose only the most important sensory characteristics and to ensure that all judges agree on the definitions of those characteristics. The number of descriptors should be limited to a maximum of 6 characteristics, especially if more than 3 samples are being presented in a session. The number of samples to present depends upon the complexity of the commodity and variables being tested, the experience of the judges, the number of characteristics to rate, etc. Again, it is best to start the panel with just a few descriptive terms and samples and then gradually increase the number of samples as the judges become more confident and accustomed to the sensory procedures. For most fruit and vegetable tests it is wise to present no more than 8 samples, since too much information may be lost when judges become fatigued.

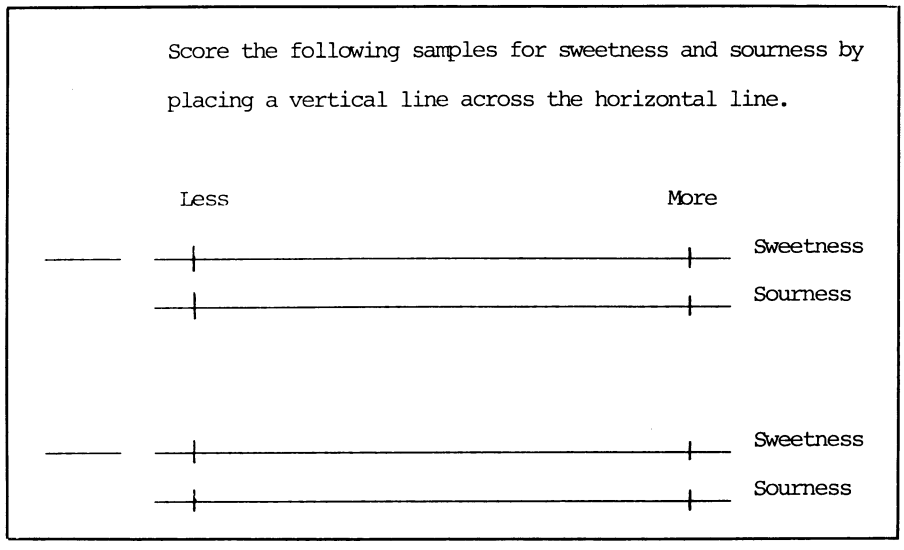

Fig. 6. Example of a descriptive scoring test score card.

Samples should be presented in random order to minimize any effects due to order of presentation.

In addition to instructions, definitions, and scoresheets, the experimenter should also prepare a calendar of treatments to be presented, a panel membership list to use as an attendance record, the list of daily treatments and the randomized order in which they are presented to each judge (decoding sheet), and record sheets for recording the data. A checklist to help prepare for a sensory panel is provided in Table 1.

Statistical methods

Choosing complete data sets (all judges, all tests, and all replications) allows for easier statistical analysis. For instance, if a judge misses a day of testing it is easier to eliminate all of his or her data from that series of tests than to deal with the missing data. Again, it is important that judges come to the panel every day. Because each judge has his or her own criteria for making decisions, judges cannot be substituted in an effort to "fill in" missing data. Three replications of judge responses are the norm. Analysis of a judge's performance over several replications of the same treatment can be used to choose the more consistent judges.

If difference testing was performed, statistical analysis merely requires counting correct responses and referring to an appropriate table (5). However, results from a descriptive test require more sophisticated handling. Analysis of variance (ANOVAR) is the most common statistical program used in evaluating results of descriptive tests. ANOVAR provides a means to study variation due to the judges, treatments, and replications. Much

Table 1. Panel preparation checklist

1. Initial steps
   a. Reserve sensory laboratory
   b. Do preliminary sample testing
   c. Choose sensory terms
   d. Recruit judges
2. Sample presentation
   a. Prepare samples (clean, peel, shell, equilibrate moisture, etc.) for training and subsequent testing
   b. Collect serving materials: ceramic cups, plates, or paper products, lids, labels, spoons, knives, spittoons, water glases, pencils
   c. Purchase panel treats
3. Paper work
   a. Prepare and duplicate—scoresheets, data sheets, instructions, definitions, and panel reminders
   b. Prepare judge attendance sheet, treatment presentation calendar, sample randomization list, and random number labels
4. Laboratory set-up: Clean booths, post instructions and definitions
5. Other steps: Prepare for corresponding chemical and physical analyses

Downloaded from https://prime-pdf-watermark.prime-prod.pubfactory.com/ at 2025-07-05 via Open Access. This is an open access article distributed under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/). https://creativecommons.org/licenses/by-nc-nd/4.0/

variation should be expected among judges and judge-treatment interactions, again because no two human subjects make decisions or react to the different treatments similarly. Variation due to replications should not be significant. If replications do vary, inconsistent judges could be eliminated. Variation in replications may also indicate insufficient judge training. When significant, treatment variance can be further defined by using the method of Least Significant Difference (LSD) or other statistical mean separation methods (5, 7).

As the ANOVAR program assumes normal distribution of data points, normalization of data prior to ANOVAR tests is sometimes performed when judges are not comfortable using the extreme values of the scoring system. We have found that if judges are presented a horizontal line that has been extended beyond the labeled extremes of the line (Fig. 6), they will tend to use a broader portion of the line and normalization may be unnecessary. Some of the assumptions and difficulties with common statistics for sensory analysis were recently discussed by O'Mahony (14).

Following ANOVAR tests, regression programs generating correlation coefficients can be used to define the interrelationships among the sensory, chemical, and objective results. Several multivariate techniques including cluster analysis, factor analysis, and principle component analysis are available to help simplify large amounts of data for ease of interpretation.

Reporting results of sensory tests

Preparation of reports and manuscripts which include sensory evaluation data requires special attention to numerous details (9, 10, 15). The following outline is proposed as a guide for thorough reporting of sensory data:

1. Statement of Purpose
   a. Type and timing of chemical, objective, and sensory tests performed
   b. Instruments and experimental methods used for objective and chemical tests
   c. Objective of sensory test (in relation to other tests performed)
2. Handling and Sampling Procedures
   a. Location and date of harvest
   b. Maturity stage at harvest
   c. Transportation
   d. Storage conditions and duration
3. Specific Sensory Evaluation Methods
   a. Number of judges
   b. Selection and training of judges
   c. Sample preparation methods (physical nature of the sample, sample size)
   d. Sample presentation (use of a reference, coding, number of samples and replications)
   e. Testing environment (individual booths, lighting)
   f. Type of sensory test (difference or descritpive tests, definitions of descriptive terms, specific instructions, scoring system)
   g. Methods used for statistical analysis

Literature Cited

1. American Society for Testing & Materials. 1968. Manual of sensory testing methods. STP 434. Amer. Soc. Testing & Materials, Philadelphia.
2. American Society for Testing & Materials. 1973. Standard recommended practice for establishing conditions for laboratory sensory evaluation of foods and beverages. E480, Part 46. Amer. Soc. Testing & Materials, Philadelphia.
3. American Society for Testing & Materials. 1978. Standard definitions of terms relating to sensory evaluation of materials and products. E253. Amer. Soc. Testing & Materials, Philadelphia.
4. American Society for Testing & Materials. 1981. Guidelines for the selection and training of sensory panel mambers. STP 758. Amer. Soc. Testing & Materials, Philadelphia.
5. Amerine, M. A., R. M. Pangborn, and E. B. Roessler. 1965. Principles of Sensory Evaluation of Foods. Academic Press, New York.
6. Cloninger, M. R., R. E. Baldwin, and G. F. Krause. 1976. Analysis of sensory rating scales. J. Food Sci. 41:1225–1228.
7. Cochran, W. G. and G. M. Cox. 1957. Experimental design, 2nd ed., Wiley, New York.
8. Institute of Food Technologists. 1964. Sensory testing guide for panel evaluation of foods and beverages. Committee on Sensory Evaluation, Inst. Food Technologists. Food Tech. 18(8):25–31.
9. Institute of Food Technologists. 1981. Guidelines for the preparation and review of papers reporting sensory evaluation data. Sensory Evaluation Div., Inst. Food Technologists. Food Tech. 35(4):16–17.
10. Institute of Food Technologists. 1981. Sensory evaluation guide for testing food and beverage products. Sensory Evaluation Div., Inst. Food Technologists. Food Tech. 35(11):50–59.
11. Kader, A. A., C. M. Heintz, and A. Chordas. 1982. Postharvest quality of fresh and canned clingstone peaches as influenced by genotypes and maturity at harvest. J. Amer. Soc. Hort. Sci. 107:947–951.
12. Kader, A. A., L. L. Morris, M. A. Stevens, and M. Albright-Holton. 1978. Composition and flavor quality of fresh market tomatoes as influenced by some postharvest handling procedures. J. Amer. Soc. Hort. Sci. 103:6–13.
13. Larmond, E. 1977. Laboratory methods for sensory evaluation of food. Pub. 1637, rev. ed. Food Res. Inst., Canada Dept. Agr., Ottawa, Ontario.
14. O'Mahony, M. 1982. Some assumptions and difficulties with common statistics for sensory analysis. Food Tech. 36(11):75–82.
15. Prell, P. A. 1976. Preparation of reports and manuscripts which include sensory evaluation data. Food Tech. 30(11):40–48.
16. Stahl, W. H. and M. A. Einstein. 1973. Sensory testing methods, p. 608–644. In: F. D. Snell & L. S. Ettre (eds.). Encyclopedia of industrial chemical analysis. Wiley, New York.
17. Stevens, M. A. and M. Albright. 1980. An approach to sensory evaluation of horticultural commodities. HortScience 15:48–50.
18. Stevens, M. A., A. A. Kader, M. Albright-Holton, and M. Algazi. 1977. Genotypic variation for flavor and composition in fresh market tomatoes. J. Amer. Soc. Hort. Sci. 102:680–689.
19. Watada, A. E., J. A. Abbott, and R. E. Hardenburg. 1980. Sensory characteristics of apple fruit. J. Amer. Soc. Hort. Sci. 105:371–375.

Downloaded from https://prime-pdf-watermark.prime-prod.pubfactory.com/ at 2025-07-05 via Open Access. This is an open access article distributed under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/). https://creativecommons.org/licenses/by-nc-nd/4.0/