



# Standard Guide for Sensory Evaluation Methods to Determine the Sensory Shelf Life of Consumer Products<sup>1</sup>

This standard is issued under the fixed designation E2454; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon ( $\epsilon$ ) indicates an editorial change since the last revision or reapproval.

## 1. Scope

1.1 This guide provides appropriate sensory testing approaches and possible decision criteria for establishing sensory shelf life of consumer products. It describes research considerations including: product selection and handling, appropriate application of various sensory test methods, selection of test intervals, and data analysis techniques for the determination of a product's shelf life end-point. As such, this document covers shelf life studies designed to identify the sensory end-point of a product's life to manage business risk and meet business needs. This guide will focus on the practical considerations and approaches, risks, and criteria that must be considered in designing, executing, and interpreting shelf life results.

1.2 Sensory shelf life is the time period during which the products' sensory characteristics and performance are as intended by the manufacturer. The product is consumable or usable during this period, providing the end-user with the intended sensory characteristics, performance, and benefits. After this period, however, the product has characteristics or attributes that are not as intended, or it does not perform the same functions as fresh products or those selected before the end of shelf life.

1.3 The goal of all shelf life determination is to estimate the time at which a consumer product is no longer usable, unfit for consumption, or no longer has the intended sensory characteristics. The determination of a sensory end-point is a function of the criteria selected, the test method used, and sampling risk. The three following test methods are most commonly used: (1) discrimination, (2) descriptive, and (3) affective. Researchers have to select criteria and methods that best suit the business risks associated with the selection of a final shelf life end point.

1.4 Research techniques designed to identify the causes of sensory shelf life changes or to develop predictive models of shelf life are beyond the scope of this document. These include the application of regression analysis to develop statistical

models designed to predict sensory shelf life, studies that assess the impact of various storage conditions, packaging materials, or product formulations on the shelf life of products, and studies designed to identify the causes of changes in product attributes over time. However, many of the research methods, experimental design considerations, and data analysis techniques discussed in this document can be applied to these other types of shelf life-related research.

1.5 This guide is not intended to provide a detailed description of how to conduct reliable sensory testing. It assumes knowledge of basic sensory and statistical analysis techniques, focusing instead on special considerations for the specific application of sensory testing method to shelf life determination.

## 2. Referenced Documents

2.1 *ASTM Standards*:<sup>2</sup>

E253 [Terminology Relating to Sensory Evaluation of Materials and Products](#)

2.2 *ASTM Manuals*:<sup>3</sup>

MNL 13 [Descriptive Analysis Testing](#)

MNL 26 [Sensory Testing Methods](#)

MNL 30 [Relating Consumer, Descriptive, and Laboratory Data to Better Understand Consumer Responses](#)

## 3. Terminology

3.1 *Definitions of Terms Specific to This Standard:*

3.1.1 *accelerated storage/aging*—subjecting a product to extreme or stressed conditions, such as elevated temperatures or humidity, to create changes in product characteristics that are assumed to be related to aging the product.

3.1.2 *control parameters for sensory shelf life determination (known as “control”)*—specific product or data set (based on previous sensory research) designated as the one to which the data from stored products are to be compared (see Section 6).

<sup>1</sup> This guide is under the jurisdiction of ASTM Committee E18 on Sensory Evaluation and is the direct responsibility of Subcommittee E18.05 on Sensory Applications—General.

Current edition approved Aug. 1, 2011. Published August 2011. Originally approved in 2005. Last previous edition approved in 2005 as E2454–05. DOI: 10.1520/E2454-05R11.

<sup>2</sup> For referenced ASTM standards, visit the ASTM website, [www.astm.org](http://www.astm.org), or contact ASTM Customer Service at [service@astm.org](mailto:service@astm.org). For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

<sup>3</sup> Available from ASTM International Headquarters, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428–2959.

3.1.3 *controlled conditions*—set of environmental parameters (temperature, humidity, light, oxygen, and so forth) that are maintained and monitored so that changes in product attributes can be evaluated relative to these parameters. For example, controlled ambient refers to maintaining a temperature of 21 to 24°C in storing the product.

3.1.4 *end point*—point at which a product no longer meets predetermined criteria as defined by test data (for example, discrimination, descriptive, or affective, or a combination thereof).

3.1.5 *pull out date(s)*—predetermined point(s) in time at which the product is removed from storage for evaluation.

3.1.6 *sensory characteristics*—any attributes of the products that are assessed using the sensory methods that measure the human response to that product.

3.1.7 *shelf life*—time period that a product may be stored before reaching its end point.

3.1.8 *shelf life testing*—method(s) to determine the effects of aging or storage conditions, or both, on product(s) characteristics for purposes of determining a product's shelf life. This testing is designed to manage distribution system risk.

3.1.9 *uncontrolled ambient*—also known as room temperature, uncontrolled conditions (that is, temperature of storage location or environmental factors, or both) which fluctuate with changes in weather, time of day, location, and so forth (see 3.1.3).

3.1.10 *zero time point*—time when the shelf life testing begins. It can be when the product is manufactured, when the ingredients equilibrate, when the product is put into storage conditions, or when the consumer is first likely to see it.

## 4. Summary of Guide

4.1 This guide is intended to cover the basic issues and practical requirements of conducting a shelf life study designed to identify end points in a product's shelf life.

4.2 *Define Research Objective*—The purpose of a shelf life project should be clearly stated before the study is implemented. Common objectives are as follows: “determine the amount of elapsed time a product remains acceptable to consumers,” or “determine specific use-by dates,” or “understand the amount of time elapsed before a sensory defect is detectable.”

4.3 *Identify Decision Risk*—Decision risk is defined as the balance of positive and negative outcomes associated with the selection of a specific end-point. Positive outcomes can be defined as correctly estimating the time a product remains within its intended sensory criteria. Negative outcomes can be defined as leaving a product on the shelf after it is acceptable, thus increasing the chances that end-users will be dissatisfied and no longer purchase the product in the future. Another type of negative outcome is to identify a too-early end-point, thus requiring shelf removal of product that is still usable to the consumer. Before the researcher embarks on a shelf life study, the risk to the consumer franchise must be balanced with the use of testing resources, as a properly designed shelf life study can use a lot of resources.

4.4 *Select Criteria to Determine the End-Point*—Products do change over time. End-point criteria can be one of the following sensory analytical or consumer criteria, or both:

4.4.1 The product's overall sensory profile has changed.

4.4.2 A product's attribute(s) (including off-notes) that is (are) known or suspected to be key to the consumers' perception of the product has changed.

4.4.3 Consumers consider the product no longer acceptable.

NOTE 1—Selection of end-point criteria depends on the extent to which the sensory attributes of the target product are required to remain unvarying and the extent to which maintaining consumer acceptance is deemed critical to the business.

4.5 *Select Sensory Test Method*—Discrimination, descriptive, or affective methods can be used to determine the shelf life of a product. Selection of the method depends on chosen end-point criteria. For example, affective testing is required if a given consumer acceptance is the chosen end-point criterion.

4.6 *Define Representative Assessors*—If discrimination or descriptive test methods are selected, the assessors used are typically trained panelists. If consumer test methods are selected, then a sample of consumers must be drawn to represent the population of potential consumers.

4.7 *Select Representative Products*—Products selected for shelf life testing must be from representative production batches and production dates and appropriately processed and packaged. In some situations, products should be subjected to typical distribution conditions (that is, vibration, temperature elevation/reduction, temperature cycling, and so forth). If testing an experimental product (for example, changes in ingredients, formulation, processing, or packaging), samples should be representative of production batches of the experimental product. The amount of product required from each production batch is dependent on the estimated length of storage, number of storage conditions, methods of evaluation, and frequency of testing.

4.8 *Determine the End-Point*—The end-point is selected based on the chosen end-point criteria, the type of product tested, the test method selected, previous knowledge of product changes over time, and the company's business needs.

## 5. Significance and Use

5.1 Measuring product changes over time serves as a practical basis for establishing the shelf life of a product. This information can also be used to assess the effects of new technology, processing, ingredients, packaging, and so forth, on the product's shelf life.

5.2 This guide provides a summary of the criteria to be considered and appropriate test methods for determining a product's sensory shelf life.

5.3 Previous sensory research with similar products, marketing research, product technology, manufacturing considerations, marketing objectives, and other business criteria can all play a part in determining sensory end point criteria.

5.4 The decision risk, end-point determination criteria, and shelf life procedure should be reviewed and agreed to by those

involved with the project, that is: R&D, Marketing, Sales, Manufacturing, Quality Assurance, Quality Control, Sensory Evaluation, and so forth.

## 6. Procedures

6.1 *Select Criteria to Determine the End-Point*—Determine the specific type of shelf life end point that will be used. There are three types of end points: (1) the product's overall sensory profile has changed; (2) a product attribute(s) that is known or suspected to be key to the consumers' perception of the product has changed; and (3) the acceptability of the product is too low. Company policy/objectives, marketplace conditions, business considerations, and risks all contribute to determining the type of shelf life end-point.

6.2 *Identify the Test Method to be Used*—This is done on the basis of the chosen type of shelf life end-point criteria. Discrimination testing, descriptive, or affective testing methods are the three major test method options.

6.3 *Clearly Determine the End Point*—The end point is established as either a significant overall difference from the control or a significant change in the intensity of one or more critical product attribute(s) or a significant decrease in acceptability to a predetermined level of acceptance. The statistical criteria for measuring significance should also be included in the end point definition (that is,  $\alpha$ ,  $\beta$ , and the effect size) along with the number of panelists needed at each testing interval.

6.4 *Select Control*—Choose the type of control product that will be used for the study:

6.4.1 *Option 1—Stable Control*—This is a typical product that is held under conditions that minimize changes over time, such as frozen or refrigerated storage or modified atmosphere. If no storage condition is known to keep changes to a minimum, this type of control cannot be used. If discrimination testing is the chosen method, this type of control is required because it is needed to conduct the test at each pull out date.

6.4.2 *Option 2—Statistical Control*—This is a set of numerical values obtained from sensory testing at zero time. Sensory data can be either descriptive analysis attribute ratings or acceptance scores, depending on the chosen testing method. If discrimination testing is the chosen method, this type of control cannot be used.

6.4.3 *Option 3—Fresh Control*—If none of the above controls are feasible, a fresh control may be obtained at each pull out date. However, this type of control can only be used if the difference between separate batches is demonstrated to be minimal prior to the study.

6.5 *Select Test Product*—Choose the test product(s) for the study. The product used for the study should be representative of intended product. If a control product is used, the test products should be obtained from the same batch as the control product. If this is not feasible, various batches can be used with the added risk of variability that may obscure the differences between the control and test product. However, the test and control products should be as close in production as possible (that is, same production date, same product location, and so forth). Furthermore, at the start of the storage period, an initial sensory test should be conducted to ensure that the test and

control products are in fact not different in either overall profile, specific descriptive attributes, or acceptance, depending on the chosen testing method.

6.6 *Determine Product Amount Needed to Conduct the Test for each Evaluation Point*—Calculate the amount of product needed for each test.

6.7 *Select Storage Conditions:*

6.7.1 *Typical Storage Conditions*—Products selected to be representative of the products in general should be stored under environmental conditions that represent the typical product distribution channel, and may include variations in temperature, humidity, light, atmosphere, air pressure, and environmental cycling (freeze/thaw, elevated temperature conditions, and so forth).

6.7.2 *Extreme Storage Conditions*—Estimates of environmental extremes and time spent in the various stages of distribution can help determine appropriate conditions. Verification of test storage conditions through the use of sensors or monitors is suggested, especially if ambient storage conditions are selected.

6.7.3 *Accelerated Storage Conditions*—Accelerated tests attempt to achieve changes in product characteristics in a short period of time. Such tests can be valuable time savers if appropriately selected. However, these tests are only approximations of how a product may behave under normal storage conditions. During accelerated storage, product changes may take place, or characteristics may develop that would not be typical otherwise. Elevated temperatures or exposure to humidity or light sources may cause different enzymatic reactions (or similar reactions but to very different degrees) than would ordinarily occur under normal storage conditions. Uses of accelerated conditions are often based on untested “rules of thumb” or beliefs rather than empirical data. Use of accelerated conditions is thus recommended only when solid research has shown what changes do in fact occur. Before determining shelf life based on accelerated conditions, establish the sensory, chemical, and mathematical relationships between accelerated conditions and typical storage conditions to ensure a high degree of reliability and validity in predicting shelf life (see ASTM MNL 30).

6.8 *Determine Sampling Plan and Evaluation Points:*

6.8.1 *Determine Baseline Point*—The first step in developing a sampling plan is to establish a baseline or “zero time” point. Choice of an appropriate baseline is determined by the nature of the product(s) and by the research objective. Examples of baseline time points include:

6.8.1.1 The date the product is manufactured,

6.8.1.2 The date the product reaches the retail shelf (the youngest product consumers would purchase),

6.8.1.3 The date the product is typically purchased, and

6.8.1.4 The date the product ingredients reach equilibrium.

6.8.2 *Determine End-Point*—The second step is to develop the expected “end point” of the product's shelf life using one or more of the following criteria:

6.8.2.1 Historical data from current or similar products,

6.8.2.2 The declared shelf life of competitive products,

6.8.2.3 Advertising or label declaration requirements,



6.8.2.4 Marketing or distribution requirements, or both,

6.8.2.5 Expected effects of packaging or ingredients, or both, and

6.8.2.6 Expected shelf life based on predicted stability of the formulation.

6.8.3 *Minimum Evaluation Points*—Considering the baseline and end points as defined above, to be 100 % of the shelf life period, choose appropriate time points for evaluations. A minimum of four evaluation points is recommended, for example, 0 % (baseline), 50 %, 100 % (end point), and one or more time points a percentage beyond the endpoint, that is, 125 %. Evaluation points beyond the end point are included in the event that the product achieves the expected shelf life and there is the possibility that the shelf life could be extended.

6.8.4 *Additional Evaluation Points*—Additional evaluation points over the shelf life period are recommended. These additional evaluations should be timed at points where significant changes in the product are expected to occur. The following examples demonstrate three sampling plans that could be used for specific product applications. These examples are designed to illustrate that sampling plans should be developed and customized based on the objectives and requirements of the study, as well as the resources available to conduct the evaluations.

6.8.4.1 For a product predicted to change most early in the shelf life period, emphasis is placed on earlier evaluations: 0 %, 15 %, 30 %, 50 %, 100 %, and some percentage beyond.

6.8.4.2 For a product predicted to change later in the shelf life period, emphasis is placed on later evaluations: 0 %, 50 %, 65 %, 80 %, 100 %, and some percentage beyond.

6.8.4.3 For a new product with little, if any, prior shelf life history, more frequent evaluation points are recommended to assure that the time of significant changes in the product are captured in evaluation. A minimum of 0 %, 25 %, 50 %, 75 %, 100 %, and one or more points beyond the expected shelf life should be planned.

#### 6.9 *Determine Product Quantities:*

6.9.1 To determine the total number of products needed to complete a shelf life study, one must take into consideration the sampling plan test intervals, the storage temperature conditions desired, the experimental design, and the sensory test methodologies best able to determine end-point criteria. Product quantities must be calculated to allow for all possible combinations of testing needs, including informal tastings.

6.9.2 It is often useful to include additional test product in each storage condition in case, for example, the product is more stable than expected or unexpected changes occur that warrant more in-depth study. An average of 20 to 50 % is common practice.

6.10 *Determine the Evaluation Plan: Multi-Point versus Single-Point Evaluation Plans*—Multi-point and single-point evaluation plans are two different evaluation plans. Neither one of these two plans is superior to the other. Each has its own advantages and disadvantages. Either one of these plans may suit each specific situation.

6.10.1 *Multi-Point Evaluations*—Multi-point evaluation plans involve tracking a single production lot over time. Test samples are pulled and tested at the predetermined evaluation

points. If control samples are used, control samples are pulled and tested at the same time. See Fig. 1 for a flow chart of the steps and decision points in a multi-point design.

6.10.1.1 *Advantages of Multi-Point Evaluation Plans*—Multi-point evaluation plans have the advantages of tracking a single production lot over the shelf life period and providing an early indication of product change.

6.10.1.2 *Disadvantages of the Multi-Point Evaluation Plans*—One disadvantage of the multi-point evaluation plans is the large quantities of control and test products necessary for multiple evaluations. Another disadvantage is that multiple evaluations are resource intensive. These disadvantages can limit the time points that can be managed or the types of sensory methods that can be used. In addition, when using multi-point evaluations, there are limitations on comparing the results of independent sensory tests conducted at different time points. Some of these limitations are addressed by using a single point evaluation plan.

#### 6.10.2 *Single-Point Evaluations:*

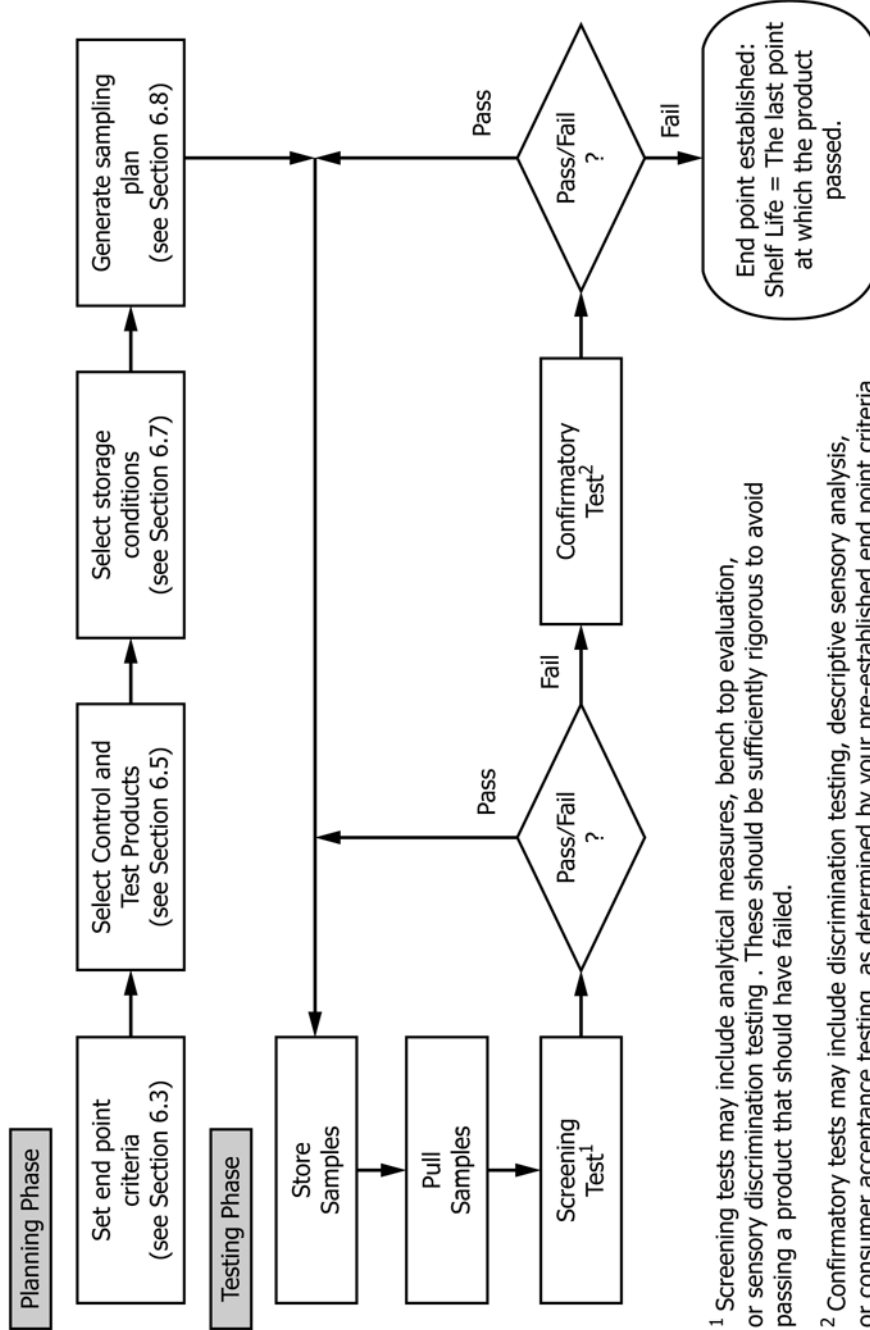
6.10.2.1 *Single-point evaluations* compare products of different ages in a single evaluation. See Fig. 2 for a flow chart showing the steps and decision points needed for single point evaluations. Products representing the selected evaluation points are accumulated over time by one of several methods. The method selected depends upon the nature of the product and the study objectives. The following are some examples:

6.10.2.2 *Staged Entry of a Single Production Lot*—A set of samples from a single production lot is placed into a control storage condition selected for its ability to preserve product characteristics. Subsets of product are removed from the control storage condition, and placed in the Test storage condition(s) at specified time points over the shelf life. Product held continuously under the control storage condition represents the zero time point control product.

6.10.2.3 *Staged Exit of a Single Production Lot*—A set of samples from a single production lot is placed into the test storage condition(s). At specified time points over the shelf life period, subsets of product are removed from the test storage condition(s) and placed into the control storage condition, to prevent any further changes in the product. Again, product held continuously under the control storage condition represents the zero time point control product.

6.10.2.4 *Staged Entry of Multiple Production Lots*—Product is collected on specified production dates and placed into the test storage condition(s). Production lots may be plant production, test batches each made from fresh ingredients, or test batches each made from the same ingredients carefully stored. The last production lot collected represents the zero time point control product.

6.10.2.5 At the evaluation point, products are pulled for a single evaluation. Depending upon available resources, products representing all of the time points may be evaluated at once, or a stepwise approach of evaluating subsets may be used. In the latter case, subsets clustered around the most likely failure point may be evaluated first. If all members of this subset fail, then a younger subset may be evaluated, or, if all members of this subset pass, then an older subset may be evaluated. Alternatively, a drill-down (that is, a progressive



<sup>1</sup> Screening tests may include analytical measures, bench top evaluation, or sensory discrimination testing. These should be sufficiently rigorous to avoid passing a product that should have failed.

<sup>2</sup> Confirmatory tests may include discrimination testing, descriptive sensory analysis, or consumer acceptance testing, as determined by your pre-established end point criteria.

FIG. 1 Multi-Point Evaluations—Process Flow and Decision Making for Shelf Life Determination

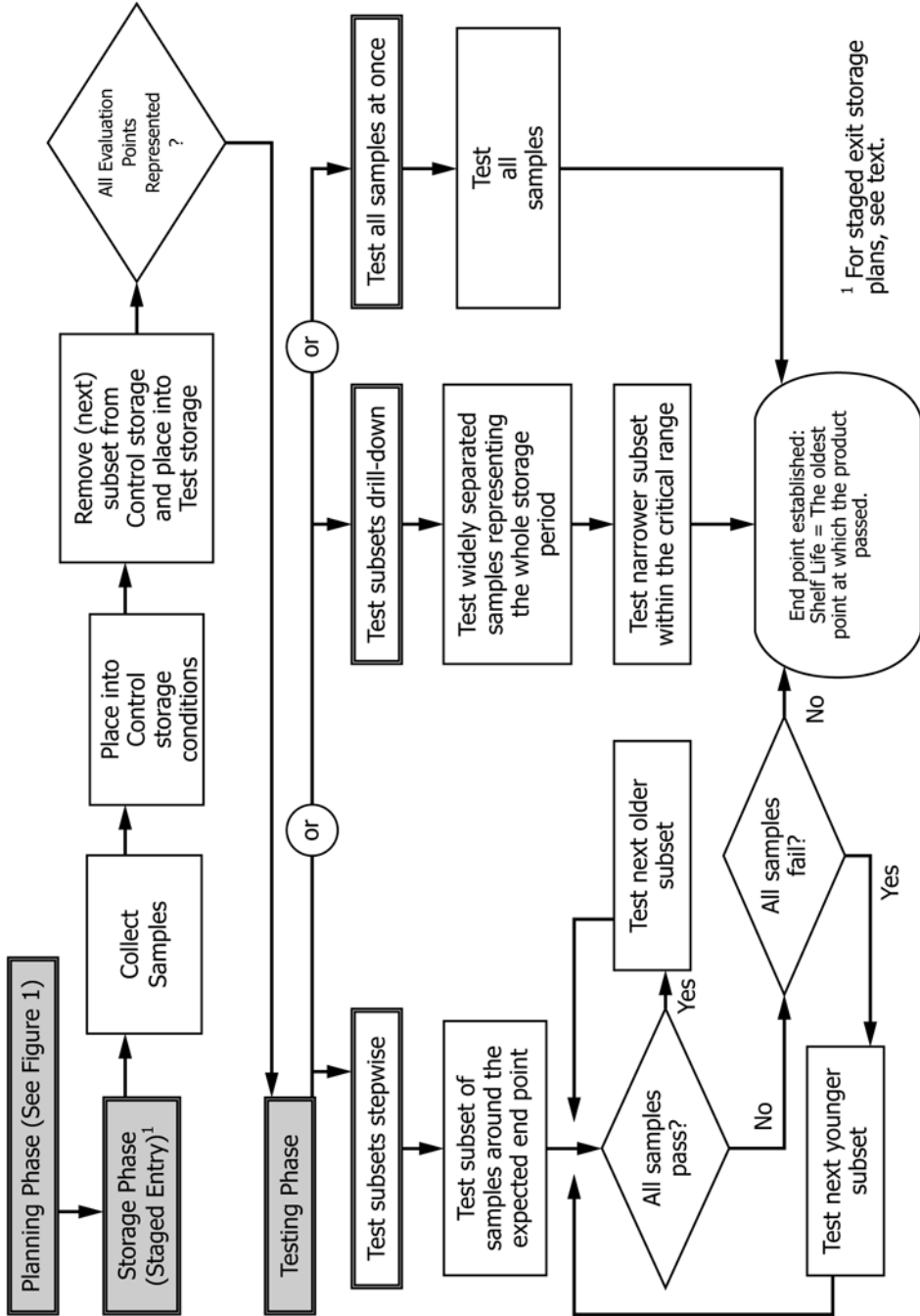


FIG. 2 Single-Point Evaluations—Process Flow and Decision Making for Shelf Life Determination

homing-in via testing the most separate products first on a defined point) approach may be employed: widely separated samples representing the whole storage period are evaluated first, allowing for a rough identification of the failure point. A narrower subset within the identified range is then tested to pinpoint the failure point.

#### 6.10.2.6 *Advantages of Single Point Evaluations:*

(1) Single-point evaluation plans minimize sources of variability due to time and panelists.

(2) Utilizing a stepwise or drill-down testing plan reduces the demand on testing resources, and thus may increase the feasibility of conducting more costly consumer testing.

(3) Stepwise or drill-down plans reduce the likelihood of missing the critical point by allowing for the collection of many more time points than will ultimately need to be tested.

#### 6.10.2.7 *Disadvantages of Single-Point Evaluations:*

(1) While single-point evaluations reduce certain sources of testing variability, they may introduce sources of product variability. For instance, if the selected control storage condition does not prevent or minimize changes in the product being studied, all products will age and differences between time points may be minimized or lost. Alternatively, a control storage condition such as freezing may introduce changes in the physical properties of certain products. Multiple production lots may introduce variability due to inconsistencies in ingredients or processing, or both. The experimenter must weigh known sources of testing variability against product and ingredient stability issues in order to select the most appropriate storage and testing design.

(2) These approaches assume that the client can wait until the most aged product is collected before receiving results, and that an early read is not required to predict or warn about a shorter shelf-life than expected. The experimenter may wish to conduct an intermediate test, perhaps halfway through the storage period, to address this concern.

(3) While single-point evaluations may reduce the load on testing resources, they may increase the demands on resources required to produce product or move product into the appropriate storage condition(s).

## 7. Test Methods

7.1 Test methods should be selected based on previously determined criteria for end point (see 6.1). Three general categories of sensory test methods include discrimination, descriptive, and affective testing. Refer to ASTM [MNL 26](#) and [MNL 13](#) for a detailed description of test methods.

7.2 Discrimination testing may be appropriate if a perceptible change in the product has been selected as the end-point criteria. The researcher should be aware, however, that discrimination testing may not be appropriate for all types of products, for example, discrimination testing cannot be used where differences are observed between two control products from the same production batch. Examples of discrimination testing methods include Duo-Trio, Triangle, Degree of Difference, Difference from Control, and 2-Alternative-Forced-Choice.

7.3 Descriptive analysis is appropriate when changes in one or more key attributes define the end point. Examples of

descriptive analysis methods include Trademark Quantitative Descriptive Analysis (QDA),<sup>4</sup> Flavor Profile, Texture Profile, and Trademark Spectrum.<sup>4</sup>

7.4 Affective testing is appropriate if a decrease in acceptability has been selected as the end-point criteria. Affective testing is conducted initially to establish a baseline and then repeated at specified intervals until predetermined end-point criteria are met. Consumer panels are most commonly used to conduct affective testing. In this case, criteria are set as a specific acceptability rating or as a significantly lower or less preferred rating from a standard. In the case where limited resources are available, internal panels can be used with a pass/fail test. In this case, criteria are set as a specific percentage of “fail” ratings.

7.5 A combination of the above methods also may be appropriate. For example, if descriptive analysis indicates a significant difference between a “control” and “test” product on one or more key attributes at six months, one may elect to follow up with affective testing to determine the impact of these changes on product acceptance. This would be appropriate if the end-point criteria were based on product changes that caused a decrease in acceptability.

## 8. Data Analysis

8.1 The type of data analysis chosen must be appropriate to the research objectives, the variables in the study, and the specific sensory method chosen. Product age or storage condition, or both, are the most commonly independent variables studied in shelf life research. Sensory characteristics or product acceptance, or both, are the most commonly studied dependent variables. The change in the product as a function of age or storage conditions, or both, is the basic purpose of shelf life testing. Refer to ASTM [MNL 26](#) and [MNL 30](#) for further discussion about appropriate statistical analysis to sensory data.

8.2 When either affective or descriptive analysis testing is used, the change in the dependent variable as a function of age or storage condition, or both, can be graphed. The mean acceptance or attribute ratings of the control sample and of the test samples are graphed. If the end-point criterion is a specific acceptance or attribute value, the graph should also show this point so that it will clearly demonstrate whether the product has or has not reached the end of its shelf life.

8.3 Statistical analysis of the data is almost always done, regardless of whether or not the results are graphed. The statistical criteria included in the definition of the end point should be used for this analysis (that is, the  $\alpha$  and  $\beta$  levels). In all cases, statistically significant results demonstrate that storage or age has changed the product above and beyond other sources of product variability.

8.3.1 *Multi-Point Designs*—At each pull-out date, data are collected and analyzed. If discrimination testing is the chosen test method, the chosen discrimination test comparing the control sample to each test sample is implemented at each

<sup>4</sup> Trademarks Quantitative Descriptive Analysis (QDA) and Spectrum are proprietary descriptive analysis techniques.

pull-out date. Refer to the appropriate standard test method for the analysis of the chosen discrimination test. If descriptive analysis or affective testing is the chosen test method, the data analysis method depends on whether the control is a physical sample available at each pull-out date (as in the case of 6.4.1 and 6.4.3) or a set of numerical values (as in the case of 6.4.2). If a physical control sample is available and the same assessors provide data for the control and test samples, use a matched pair sample *t*-test. If different assessors provide the data for the control and test samples, then a two sample *t*-test should be used. If the control is a set of numerical values and there is only one test sample, use a one-sample *t*-test to assess whether the observed acceptance or attribute mean rating is significantly different from the control's numerical value.

**8.3.2 Single-Point Designs**—In the case of staged entry of multiple production lots, the data analysis cannot segregate the storage variation from the production lot variation. Hence, this method should be used only if production lot variation was demonstrated, *a priori*, to be minimal. If discrimination testing is the chosen testing method, the chosen discrimination test comparing the control sample to each test sample is implemented at the given evaluation date. If descriptive analysis or acceptance testing are the chosen testing methods, use a repeated measures analysis of variance and pair-wise post-hoc comparisons.

**8.4 Interpretation of Typical, Extreme, and Accelerated Storage Conditions**—The data analyses described above are appropriate for any of these storage conditions. However, the results must be interpreted differently. If extreme-condition or accelerated-storage samples fail, the relationship between the extreme/accelerated conditions to actual distribution must be

considered. Rules of thumb such as “each 10° above ambient temperature is equivalent to three months of storage time” are often used to interpret shelf life data. These rules of thumb should only be used when there are data showing the relationship between storage condition and aged product. It is possible that extreme conditions or accelerated storage results in sensory attributes that are not present in products in typical distribution conditions. See 6.7.3.

8.5 Based on the definition of the end point and the results from the data analysis, determine the shelf life of the product.

8.6 Decision matrices, which include test methods, results, and end point criterion for discrimination, descriptive, and affective testing are presented in Figs. 1 and 2.

8.7 A detailed explanation of survival analysis or other possible modeling techniques that could be applied to shelf life data is beyond the scope of this document. Details of such modeling techniques can be found in other references (Hough et al., 2003). Often, such studies explore the relationships between various storage conditions, packaging materials, or product formulations, to determine impact on the rate of change leading to shelf life end-point. The experimental design considerations discussed in this document apply in studies where shelf life end-points need to be determined in a descriptive and practical manner, rather than a study in which variables are formally studied and a predictive model is the ultimate objective.

## 9. Keywords

9.1 flavor stability testing; sensory changes during storage; sensory changes over time; shelf life testing

## BIBLIOGRAPHY

- (1) Berger, K. G., “Practical Applications of Accelerated Stability Test to Rancidity Problems in Food Processing,” *Food Technology*, Vol 6, No. 3, 1972.
- (2) Bishop, J. R., and White, C. H., “Assessment of Dairy Product Quality and Potential Shelf Life—A Review,” *Journal of Food Protection*, Vol 49, No. 9, 1986, pp. 739-753.
- (3) Blandt-Jonhson, R. C., and Johnson, N. L., *Survival Models and Data Analysis*, John Wiley & Sons, Inc., New York, NY, 1980.
- (4) Cecil, S. R., and Woodroof, J. G., “The Stability of Canned Foods in Long-term Storage,” *Food Technology*, Vol 17, No. 5, 1963, p. 131.
- (5) Charalambous, G., “The Shelf Life of Foods and Beverages,” *Developments on Food Science*, Vol 12, Elsevier Science Publications Co. Inc., NY, 1986.
- (6) Cox, D. B., and Oaks, D., *Analysis of Survival Data*, Chapman, London, 1984.
- (7) Dethmers, A. E., and Rodriguez, N. C., “Food Product Shelf Life: Sensory Panel Evaluations and the Open Dating Issue,” *Food Product Dev.*, 1975, pp. 96-102.
- (8) Dethmers, A. E., “Utilizing Sensory Evaluation to Determine Product Shelf Life,” *Food Technology*, September, 1979, pp. 40-43.
- (9) Gacula, M. C., Jr., “The Design of Experiments for Shelf Life Study,” *Journal of Food Science*, Vol 40, 1975, p. 399.
- (10) Gacula, M. C., Jr., and Kubala, J. J., “Statistical Models for Shelf Life Failures,” *Journal of Food Science*, Vol 40, 1975, p. 404.
- (11) Gacula, M. C., and Singh, J., *Statistical Methods in Food and Consumer Research*, Academic Press, Inc., Orlando, FL, 1984.
- (12) Griffiths, N., “Sensory Analysis in Measuring Shelf Life,” *Food, Flavours, Ingredients, Packag. Process*, Vol 7, No. 9, 1985, pp. 47-48.
- (13) Hough, G., Langohr, K., Gomez, G., and Curia, A., “Survival Analysis Applied to Sensory Shelf Life of Foods,” *Journal of Sensory and Nutritive Qualities of Foods*, Vol 68, No. 1, 2003, pp. 359-362.
- (14) Hough, G., Garitta, L., and Sanchez, R., “Determination of Consumer Acceptance Limits to Sensory Defects Using Survival Analysis Statistics,” Paper presented at Fifth Pangborn Conference, Boston, MA, July, 2003.
- (15) Hearne, J. F., “Long-Term Storage of Foods,” *Food Technology*, Vol 18, No. 3, 1964, p. 318.



- (16) Harris, N. E., Wescott, D. E., and Henick, A. S., "Rancidity in Almonds: Shelf Life Studies," *Journal of Food Science*, Vol 37, 1972, pp. 824-827.
- (17) IFT, "Shelf Life of Foods. Inst. of Food Technologists Expert Panel on Food Safety and Nutrition," *Journal of Food Science*, Vol 39, 1974, p. 861.
- (18) Jul, M., "Quality and Stability of Frozen Meats," *Quality and Stability of Frozen Food*, W. B. Van Aradel, M. J. Copley, and R. L. Olson, eds., Wiley Interscience, NY, 1969.
- (19) Kwoleck, W. F., and Bookwalter, G. N., "Predicting Storage from Time-Temperature Data," *Food Technology*, Vol 25, No. 10, 1971, pp. 51-63.
- (20) Labuza, T. P., and Schmidl, M. K., "Accelerated Shelf Life Testing of Foods," *Food Technology*, Vol 39, No. 9, 1985, pp. 57-62, 64, 134.
- (21) Labuza, T. P., and Schmidl, M. K., "Use of Sensory Data in the Shelf Life Testing of Foods. Principles and Graphical Methods for Evaluation," *Cereal Foods World*, Vol 33, No. 2, 1988, pp. 193-206.
- (22) Lawless, J. E., *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, Inc., New York, NY, 1982.
- (23) Maxcy, R. B., and Wallen, S. E., "Heterogeneity of Samples as a Problem in Shelf Life Prediction," *Journal of Food Protect*, Vol 46, No. 6, 1983, pp. 542-544.
- (24) Meilgaard, M. C., Civille, G. V., and Carr, B. T., *Sensory Evaluation Techniques*, 2nd Ed., CRC Press, Boca Raton, FL, 1991, pp. 198-199, 222-225.
- (25) Peryam, D. R., "Consumer Preference Evaluation of the Storage Stability of Foods," *Food Technology*, Vol 18, No. 9, 1964, pp. 214-217.
- (26) Powers, J. J., "Using General Statistical Programs to Evaluate Sensory Data," *Food Technology*, June, 1984, pp. 74-84.
- (27) Ragnarsson, J. O., and Labuza, T. P., "Accelerated Shelf Life Testing for Oxidative Rancidity in Foods—A Review," *Food Chemistry*, Vol 2, No. 4, 1977, pp. 291-308.
- (28) Stone H., and Sidel, J. L., *Sensory Evaluation Practices*, 2nd edition, Academic Press, Orlando, FL, 1993.
- (29) Van Arsdel, W. B., "The Time-Temperature Tolerance of Frozen Foods. Introduction—The Problem and the Attack," *Food Technology*, Vol 2, No. 1, 1957, pp. 28-34.
- (30) Walezko, P., and Labuza, T. P., "Accelerated Shelf Life Testing of an Intermediate Moisture Food in Air and in Oxygen Free Atmosphere," *Journal of Food Science*, Vol 41, No. 6, 1976, pp. 1338-1344.

*ASTM International takes no position respecting the validity of any patent rights asserted in connection with any item mentioned in this standard. Users of this standard are expressly advised that determination of the validity of any such patent rights, and the risk of infringement of such rights, are entirely their own responsibility.*

*This standard is subject to revision at any time by the responsible technical committee and must be reviewed every five years and if not revised, either reapproved or withdrawn. Your comments are invited either for revision of this standard or for additional standards and should be addressed to ASTM International Headquarters. Your comments will receive careful consideration at a meeting of the responsible technical committee, which you may attend. If you feel that your comments have not received a fair hearing you should make your views known to the ASTM Committee on Standards, at the address shown below.*

*This standard is copyrighted by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959, United States. Individual reprints (single or multiple copies) of this standard may be obtained by contacting ASTM at the above address or at 610-832-9585 (phone), 610-832-9555 (fax), or service@astm.org (e-mail); or through the ASTM website (www.astm.org). Permission rights to photocopy the standard may also be secured from the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, Tel: (978) 646-2600; http://www.copyright.com/*