Designation: ~~E2454 – 05 (Reapproved 2011)~~ E2454 – 19

Standard Guide for Sensory Evaluation Methods to Determine the Sensory Shelf Life of Consumer Products[1]

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 (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This guide provides ~~appropriate~~recommended sensory testing approaches and ~~possible~~ decision criteria for establishing the sensory shelf life of ~~consumer products.~~ consumed products, including food, personal care, and household products to manage business risk. It describes research considerations including: product selection and handling, appropriate application of ~~various~~specific sensory test methods, selection of test intervals, and data analysis techniques for the determination of a product's sensory 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 sensory 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.~~

1.2 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.

1.3 The shelf life measures in this guide refer to foods, household and personal care products stored as the manufacturer intended and do not account for changes in sensory properties occurring after opening, partial consumption or in-home storage. Once products have been manufactured, packaged and sent through the distribution channels, the condition of the products is not typically under study. However, a company may wish to include such variables in their shelf life studies when there is a need to evaluate the sensory quality of their products as they go through distribution channels and/or in-home storage and use.

1.4 This guide is not intended to address non-sensory issues related to the shelf life of food, including microbial contamination and chemical changes of products associated with aging, nor is it intended to address potential safety issues associated with aging food and non-food consumer products.

1.5 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

---

[1] 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~~Feb. 1, 2019. Published ~~August 2011~~February 2019. Originally approved in 2005. Last previous edition approved in ~~2005~~2011 as ~~E2454~~E2454 – 05 (2011).~~-05.~~ DOI: ~~10.1520/E2454-05R11.~~10.1520/E2454-19.

2. Referenced Documents

2.1 *ASTM Standards:*[2]
E253 Terminology Relating to Sensory Evaluation of Materials and Products
2.2 *ASTM Manuals:*[3]
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:*
3.1.1 For definitions of terms used in this guide see Terminology E253.
3.2 *Definitions of Terms Specific to This Standard:*
3.2.1 *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.

3.2.2 *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).

3.2.3 *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.

3.2.4 *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).

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

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

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

3.2.8 *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.

3.2.9 *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~~3.2.3).

3.2.10 *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.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.
[3] Available from ASTM International Headquarters, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428–2959.

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, α, β, 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 overage 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 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.

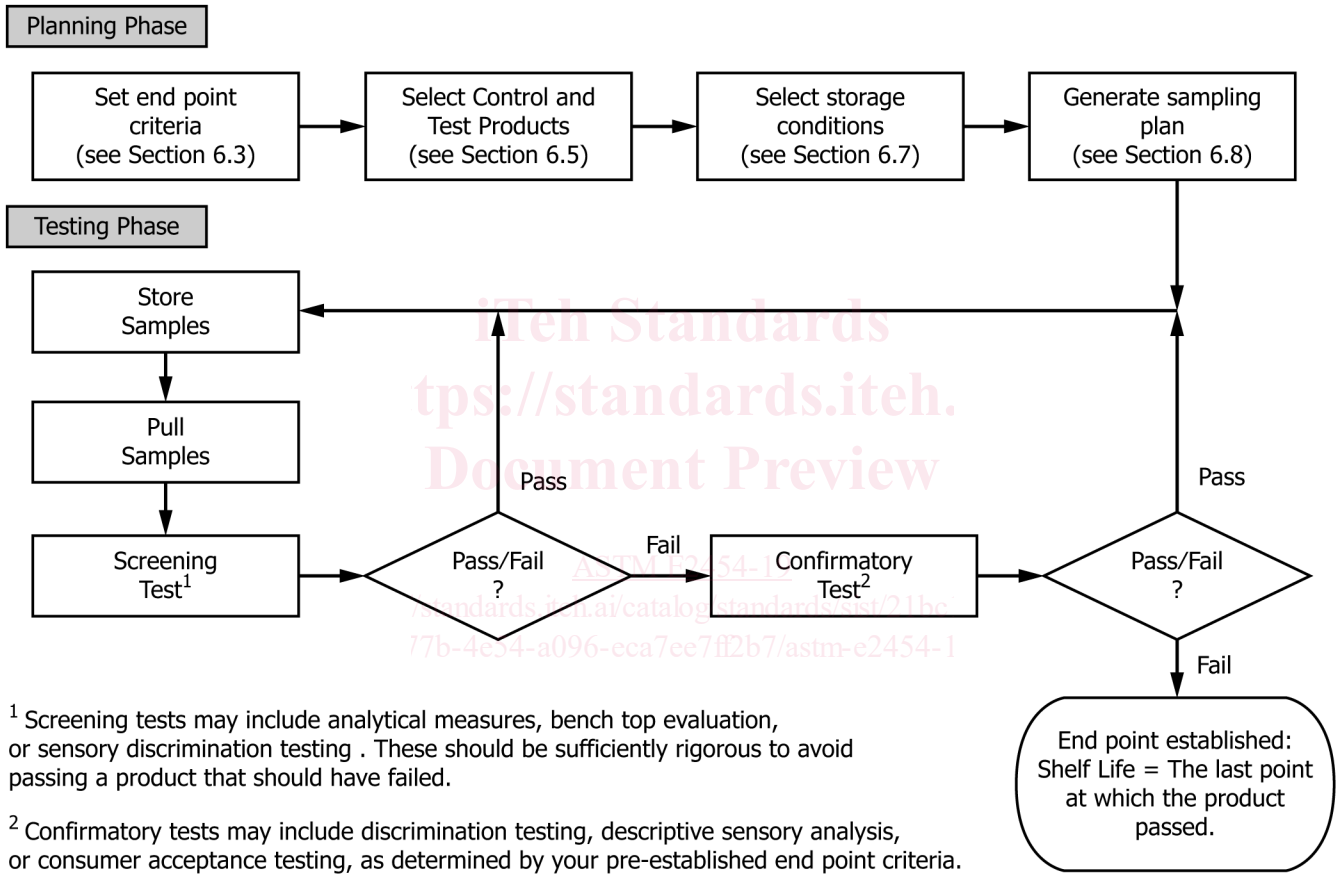
(2) Utilizing a stepwise or drill-down testing plan reduces the demand on testing resources, and thus may increase the feasibility of conducting more costly consumer testing.

(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.



[1] 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.

[2] 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