

Introduction to Sensory Evaluation

I. Introduction and Objective	1
II. Historical Background	7
III. Development of Sensory Evaluation	11
IV. Defining Sensory Evaluation	13
V. A Physiological and Psychological Perspective	16

I. Introduction and Objective

Since publication of the second edition of this book a decade ago, there have been many developments in the science and the application of sensory evaluation that directly or indirectly have had an impact. Today, just about all consumer products companies in the food and beverage industry as well as other industries, for example, home care and personal care industries, are aware of sensory evaluation and most agree that it has a role within their company. Marketing research and brand management professionals also are giving increased recognition to sensory information. Such recognition has yielded benefits for the profession in the form of improved status (and increased compensation), and for some, a bigger/louder voice in the product decision-making process. It is the latter which has the greater effect in the longer term. These developments have also resulted in more support for research and more course offerings at the University level. However, some fundamental, as well as some practical, issues remain to be considered or re-considered. New professionals tend to rely on pre-packed software that provides not only data capture capabilities, but also design and analyses options. While such capabilities are a significant enhancement, many also provide designs that are neither balanced properly nor relevant for anything other than a standard test. Reliance on such packages make for an easy entry for the inexperienced professional but have the potential for misapplication when the action taken is to modify the problem to fit the program. There has been a decline in understanding and appreciation for the consequences of not using qualified subjects, a tendency to limit replication (in sensory analytical tests) for cost savings purposes,

2 Introduction to Sensory Evaluation

and as already noted, use of statistical packages with default systems that yield results that on the surface make sense but have no real basis or there is no awareness of their weaknesses. We will explore these issues in more detail later in this book.

Using sensory information as a part of a marketing decision has given it unprecedented attention; being able to identify and quantitatively model the key drivers for a product's acceptance is now generally recognized as a core resource for any sensory program. It is acknowledged to be a powerful approach to optimizing product preference; however, this has only been possible with use of descriptive analysis to identify the specific sensory characteristics. The next logical step in this process has been to incorporate this information with other cognitive measures such as imagery. Exploiting this information to the fullest extent possible has enabled companies to grow their market share as well as implement cost savings through better use of technology, etc. All this has been possible as a direct result of use of sensory resources, better understanding about the measurement of human behavior, combined with a more systematic and professional approach to the testing process. Much of this progress has been achieved within the technical and marketing structures of those organizations that recognized the unique contributions of sensory evaluation. In the past, these activities were the exception, today it is a more common occurrence, again reflecting the increased awareness of sensory evaluation by those in and outside the field. For a summary of these developments, the reader is directed to Schutz (1998). However, much more needs to be done, in part because the link between sensory, marketing, and production is not strong, and in part because there is a lack of appreciation for the principles on which the science is based. For some, sensory evaluation is not considered as a science capable of providing reliable and valid information. This is a not so uncommon "perception," fostered in part by the seemingly simplistic notion that anyone can provide a sensory judgment. We are born with our senses and barring some genetic defects, we are all capable of seeing, smelling, tasting, etc. It certainly seems simple enough, so why should a technologist or a brand manager believe results from a test that are inconsistent with their expectations and their own evaluations? Alternatively, product experts and reporters such as those responsible for wine and food reviews in the public press have a significant impact on the success of products and businesses based on their reviews which purport to be based on the senses and by default, sensory evaluation. Myths are created and perpetuated based on hearsay simply by being in a position of authority. Suffice to say, not all of what one reads should be believed. We will have more to say about this later in this chapter and in the chapters on discrimination and descriptive analysis. As a result, demonstrating that there is a scientific basis to the discipline continues to be a challenge because the basic principles of perception are being lost in favor of quick solutions or a lack of time to do a test correctly. In the case of the latter, the authors have experienced the situation in which time restrictions take precedence over using an appropriate method even though there was a high risk of obtaining inadequate information using a method incorrectly! It is no longer a surprise to hear statements such as, "We don't have the time or money to do it right, but we will be able to do it over again later." It takes a lot of effort to overcome this kind of thinking. Since the previous editions of this book, advances continue to be made, albeit at a slow pace,

not because test procedures are inadequate, but as noted previously, the science is not readily acknowledged as such. In all fairness, it should be mentioned that sensory professionals have not been effective spokespersons for their work or for the science. In one company, sensory evaluation will be used successfully, but in another it will be misused or the information will be ignored because it is inconsistent with expectation. Unfortunately, this latter situation has encouraged use of other information sources or to develop competing test capabilities in the hope of obtaining acceptable information without fully appreciating the consequences.

Over the years, numerous efforts have been made and continue to be made to develop a more permanent role for sensory evaluation within a company. Reviewing the technical and trade literature shows that progress in the development and use of sensory resources continues. There has been a noticeable increase, and much of the impetus continues to come from selected sectors of the economy, notably foods and beverages and their suppliers (Piggot, 1988; Meiselman and MacFie, 1996; Lawless and Heymann, 1999; Jackson, 2002). In their seminal textbook on sensory evaluation published almost four decades ago, Amerine *et al.* (1965) correctly called attention to three key issues: the importance of flavor to the acceptance of foods and other products, the use of flavor-related words in advertising, and the extent to which everyday use of the senses was largely unappreciated, at that time. Perhaps a secondary benefit of today's concerns about food safety has been an awareness by consumers of the sensory aspects of the foods they purchase.

It is apparent that current interest in sensory evaluation reflects a more basic concern than simply being able to claim use of sound sensory evaluation methodologies. In a paper published in 1977, Brandt and Arnold described the results of a survey on the uses of sensory tests by food product development groups. Their survey provided insight into some of the basic issues facing sensory evaluation. While the survey is dated, the information remains relevant and much of it continues to be confirmed based on more recent surveys fielded by the Sensory Evaluation Division of the Institute of Food Technologists (see below). The results were especially notable for the extent (or lack of) of the awareness of sensory evaluation by the respondents. Of the sixty-two companies contacted, fifty-six responded that they were utilizing sensory evaluation. However, descriptions of tests being used revealed that confusion existed about the various methods; for example, it was found that the triangle test (a type of discrimination test) was the most popular, followed by hedonic scaling (a type of acceptance test) and paired comparison (either an acceptance test or a discrimination test). Since these methods and the other methods mentioned in the survey provide different kinds of information, it is not possible to evaluate the listing other than to comment that the groups in most companies contacted appeared to be using a variety of procedures.

Also, there was confusion about accepted test methods and the information that each provides. For example, single-sample presentation is not a test method, yet twenty-five of the fifty-six companies responded that it was one of the test methods in use. It is, in fact, a serving procedure and such responses may reflect poorly worded or misunderstood questions in the survey. Another example of the confusion is "scoring," which was said to be in use by only seven of the fifty-six companies contacted. However, all sensory tests entail some kind of scoring. Statistical tests

4 Introduction to Sensory Evaluation

included in the survey elicited similarly confusing responses. The failure to define the terms confirms that the food industry utilizes sensory evaluation but there is confusion as to what methods are used and for what applications. As previously noted, sensory evaluation as a science is not well understood, methods are not always used in appropriate ways, and results are easily misused. While there continues to be a lack of qualified professionals, more are being trained. More books on sensory evaluation continue to be published; however, the science still has not achieved a status commensurate with its potential. As noted by Stone (1999), separating science from mythology is a challenge that is still with us. Similarly, its role within individual companies is far from clear. One of the goals of this book is to provide a perspective on all these issues and how one should develop resources and integrate them into the business environment. This is not to imply that the scientific basis of sensory evaluation is overlooked. In fact, considerable attention is given to the science of sensory evaluation because without it, any use of sensory information will be seriously compromised.

A focus of this book is on a systematic approach to the organization, development, and operation of a sensory program in a business environment. While emphasis is directed toward the more practical aspects of sensory evaluation and approaches to implementing a respected and credible program, attention also is given to more fundamental, underlying issues, including a comparison of specific test methods, experimental design, the reliability and validity of results, and related topics. From a sensory evaluation perspective, the twin issues of reliability and validity are of paramount importance, and are integral to developing a credible program and providing actionable recommendations within the context of that company's business and brand strategy (see Aaker, 1996). From a business perspective, it is these latter issues that loom as most important along with knowing which product formulation best meets the objective. Providing actionable recommendations is critical as it builds trust that others will place on those recommendations and the extent to which managers will act on those recommendations. The idea that volumes of data will convince one's peers and superiors is not an approach that is viable, particularly in the world of business. Sensory professionals must communicate information clearly to ensure that superiors, peers, and subordinates understand what was done and what the results mean in terms of meeting that test's specific objective.

It should be clear to the reader as to the importance of the relationship between reliability and validity of results and the credibility assigned to sensory recommendations, and in a global sense, the credibility of a sensory program itself. While it is logical to assume that a test result should speak for itself, and that quality information will be heard, understood, and acted on accordingly; in fact, the opposite can and does occur a sufficient number of times to directly impact on the effective utilization of sensory evaluation. It raises important questions as to why sensory information is not better understood and, therefore, better used. Whether this situation reflects a basic lack of understanding of results from a test and/or a lack of understanding of the role of sensory evaluation in business, or a failure of sensory to understand the background to a specific request is not clear. It would appear that there is no single or simple answer. However, this situation can be better understood

if one considers the nature of the sensory evaluation process. Product evaluation is a multi-step process in which a group of individuals respond to stimuli (a set of products) by marking a scorecard according to a specific set of instructions. These individuals are participating in this test because of their demonstrated sensory skills with that particular category of products. The responses are usually marks on the scorecard or a similar designation on a computer screen. The marks are then converted to numerical values for computation. Several aspects of this process warrant comment – the first is whether the test plan is appropriate for the problem, the second is whether the scorecard is appropriate, third is how the information will be collected (numerical or marking some type of scale), the fourth relates to the type of scale used, and the fifth is the type(s) and relevance of analysis used. Having a good understanding of why a test is being requested represents the first and perhaps most important task for a sensory professional. Failure to obtain such information is a clear sign of impending problems, beginning with establishing an objective to determining how the results will be communicated. The scorecard lists the questions and the means by which judgments will be captured, and by default, what analysis or analyses will be used. Similarly, it will determine which subjects should participate. A common problem for many is the seemingly innocuous request for subjects to provide numerical judgments because they are easy to capture (and analyze). Yet, more than 30 years ago, it was clearly demonstrated that there are number biases, some numbers have residual meanings, numbers connote a right or wrong answer, and subjects will change their use of numbers for no apparent reason, usually in the middle of a test. Scaling also warrants comment; however, there is a more detailed discussion about this topic in Chapter 3. One issue is the concept of using the same scale in all tests, another relates to using the “the standard scale” from that company (which usually has no basis in any literature) or a belief that there is a universal scale. There is no question that comparisons of results are made easier when the same scale(s) is used; however, few problems lend themselves to use of the same scale in every test. Not all questions can be answered with the same scale and this should not be a surprise. Nor should it be expected that there is one scale that is more sensitive than all other scales. Here, too, the nature of the problem and the products will help determine the most appropriate scale.

Subjects also have an impact on a program’s credibility. How they were selected and what kind of training, if any, they received are important considerations as are their responses, in terms of both their sensitivity and their reliability. To a degree, a panel of subjects can be considered as functioning like a null instrument, recording what is perceived. The implication of a human (or group of humans) functioning like an instrument has obvious appeal in a technical sense, particularly to those who mistakenly envision an invariant system yielding the same numerical values time and time again. In fact, this latter concept has considerable appeal, particularly among those with a poor understanding of the perceptual process. The realities of the situation are, of course, quite different. Subjects, no matter what their level of skill or number of years of training and practice, exhibit differences in sensitivity from one another, and differences in variability that is unique (to that individual). Some training programs imply (e.g. see Spectrum Analysis in Meilgaard *et al.*, 1999) that

6 Introduction to Sensory Evaluation

this sensitivity and accompanying variability can be overcome through training and use of appropriate references. Such training, as much as 10 hours per week and often lasting 4 or more months, has considerable appeal (it is so special, it must be correct); however, independent evidence of the success of such an approach is not readily demonstrated nor is it consistent with our knowledge of human perception and the physiology of the senses. Such an approach is a form of behavior modification rather than a means of capturing responses as a function of a stimulus (whether that stimulus is a purified chemical or a consumer ready beverage). This ignores the fact that changing one's response does not necessarily mean that one's perceptions also are changed, nor does it recognize the variability inherent in humans and in products. In each instance, the end result is to overcome what appear to be limitations of the sensory process in the mistaken belief that they will make results more acceptable. Short of directly telling an individual what should be an answer, there always will be variability. Nonetheless, the requestor of a test expecting that test to yield an invariant result (the same response or numerical value every time) is disappointed and concerned about this unique information source. This disappointment also is accompanied by reservations as to any conclusions and recommendations; that is, the issue of credibility arises. Alternatively, when results are not in agreement with information obtained elsewhere (and often not sensory information) and no attempt is made to understand and explain the basis for any disagreement, then further erosion of program credibility will occur. The success of a program, and particularly its credibility, begins with having a plan and an organized effort, and making sure that the testing process follows accepted procedures and practices; what method was used, who the subjects were and how they were selected, what test design was used, how the data were analyzed, including the evidence of reliability, and so forth. In a business sense, it begins with an explanation of what sensory information is, how the results are communicated and whether they are understood, and ends with actionable recommendations being implemented to everyone's satisfaction. While these issues are discussed in detail in succeeding chapters, their inclusion here is to provide an initial perspective to the issue of business credibility and direct involvement of sensory evaluation in product decisions. Without an organized product evaluation effort and demonstrated reliable and valid results that are communicated in a way that is readily understood, one is returning to reliance on product experts, the *N* of 1, who made product decisions by virtue of their expertness and not because there were data to support their judgments. Decisions derived in this manner are neither new nor are they unusual; however, they make it very difficult for those individuals trying to organize and operate a credible sensory test program. As noted by Eggert (1989) and more recently by Stone and Sidel (1995), sensory evaluation must develop a strategy for success. It must reach out to its customers, it must educate them about the benefits that can be realized from using sensory information. At the same time, it must gain management support through an active program of selling its services and how the company can benefit from those services.

This book is not intended as an introduction to the topic of sensory evaluation. Nonetheless, for some aspects of sensory evaluation, considerable detail is provided in an almost stepwise manner. Readers, however, will probably derive more from this

book if they have a basic understanding of sensory evaluation, experimental design, and statistics, and especially the perceptual process.

Where appropriate, background information sources are cited in this text and should be considered as recommended reading. In addition to its benefit to the sensory professional, this book is intended to serve as a guide for the research and development executive seeking to have a capability in sensory evaluation and to develop a more efficient and cost effective product development program. It should be of interest also to marketing, market research, and technical managers, all of whom have an interest in their company's products and their quality as measured by consumer responses and through sales, market share, and profitability.

II. Historical Background

Of the many sectors of the consumer products industries (food and beverage, cosmetics, personal care products, fabrics and clothing, pharmaceutical, and so on), the food and beverage sectors provided much early support for and interest in sensory evaluation. During the 1940s and through the mid-1950s, sensory evaluation received additional impetus through the US Army Quartermaster Food and Container Institute, which supported research in food acceptance for the armed forces (Peryam *et al.*, 1954). It became apparent to the military that adequate nutrition, as measured by analysis of diets or preparation of elaborate menus, did not guarantee food acceptance by military personnel. The importance of flavor and the degree of acceptability for a particular product were acknowledged. Resources were allocated to studies of the problem of identifying what foods were more or less preferred as well as the more basic issue of the measurement of food acceptance. These particular problems were apparently forgotten during the 1960s and early 1970s when the federal government initiated its "War on Hunger" and "Food from the Sea" programs. The government's desire to feed the starving and malnourished met with frustration when product after product was rejected by the recipients primarily because no one bothered to determine whether the sensory properties of these products were acceptable to the targeted groups. This is not to suggest that each country's ethnic and regional food habits and taboos were not important but rather, in the context of these government programs, there was scant attention given to the sensory evaluation of the products as they were being developed. This situation continues to exist because there remains a fundamental lack of appreciation for the importance of sensory perception on food choice behavior.

The food industry, possibly taking a cue from the government's successes and failures in sensory evaluation, began to provide support for this emerging science. Although many industries have since recognized its value in formulating and evaluating products, general appreciation for sensory evaluation as a distinct function within a company remained minimal until this past decade. In general, there is agreement on the role of sensory evaluation in industry but not necessarily how sensory evaluation should be organized and how it should operate within a company.

8 Introduction to Sensory Evaluation

As with any emerging discipline, divergent opinions and philosophies on sensory evaluation exist both within and outside the field. It is not necessary that we examine all these opinions (a majority) and philosophies (a minority) in detail. However, some discussion of them is appropriate, to enable the reader to gain a greater appreciation for the problems involved in the organization and operation of a sensory program.

The food industry (as well as many other consumer products industries) traditionally viewed sensory evaluation in the context of the company “expert” (the *N* of 1) who through years of accumulated experience was able to describe company products and set standards of quality by which raw materials would be purchased and each product manufactured and marketed. Examples of such “experts” include the perfumer, flavorist, brewmaster, winemaker, and coffee and tea tasters. In the food industry, experts provided the basis for developing the “cutting sessions” and “canning bees” (Hinreiner, 1956). In the canning industry, products usually were evaluated on a daily basis and in comparison with competitive products, new products, etc. In addition, there were industry-wide cutting bees to assess general product quality. These sessions enabled management and the experts to assess product quality at their own plants, as well as to maintain a familiarity with all other company’s products. This process continues today in some companies as well as in trade associations seeking to solve common problems that are usually related to product quality. In recognizing the purpose of the cutting bee and its overall function, Hinreiner (1956) described the efforts undertaken to improve the quality of the information derived from one group, the olive industry. The Processed Products Branch of the Fruit and Vegetable Division of the United States Department of Agriculture (File Code 131A-31, 1994) has updated its guidelines for approved illumination for cutting rooms, an action that recognizes the importance of providing a more standardized environment for product evaluations. In addition to the evaluation itself, record-keeping was formalized, making it possible to compare results from 1 year with another and thus provide for a great degree of continuity. It is important to remember that the industry recognized a problem, and with assistance from sensory evaluation, took some action to improve its information about product quality. This activity continues today, especially in those industries that rely on basic agricultural products that do not experience substantial value-added processing; for example, the wine industry, processed fruits and vegetables, olive oil, etc.

In addition to cuttings, product experts also have endured; however, the impact of many of these experts has been considerably eroded, or in some instances, is exerted only indirectly. In retrospect, the results accomplished (and failed to be accomplished) by these experts and the cutting sessions were quite remarkable. By and large, experts determined which alternatives, from among many alternative ingredients, products, and so forth, were appropriate for sale to the consumer. Their success reinforced their role for establishing quality standards for particular products, such as canned fruits and vegetables, and these standards, in turn, received additional support through official USDA standards that referenced these results. As long as the food industry was involved solely in the preserving of a basic agricultural crop, for example, frozen peas, canned fruits and vegetables, or a juice, then it was relatively

easy (or uncomplicated) for the product expert to understand a particular product category and to make reasonably sound recommendations.

In the early stages of the growth of the food-processing industry and where competition was primarily regional in character, such standards and evaluative procedures by experts were extremely useful. In most industries, experts also created a variety of scorecards (and unique terminology) to serve as a basis for maintaining records and presenting a more scientific process. Subsequently, numerical values were also assigned to the scorecards, as described by Hinreiner (1956). These scores soon became targets or standards; for example, the 100-point butter scorecard, the ten-point oil quality scale, and the twenty-point wine scorecard all had specific numbers that connoted levels of product acceptance (equated with quality). All of these and others continue to be used. Certain values became fixed in people's minds, and they were transposed inappropriately into measures of consumer acceptance, creating a multitude of new problems. That some of these scorecards have survived virtually intact after 50 years is remarkable, considering their faults. Where they have not survived one can usually find the concept still alive, particularly that of the single number equated with quality and the associated belief of the invariance of the expert. While it is more common in quality control, the re-emergence of experts in sensory evaluation is somewhat surprising (and discouraging). Not only does it reflect a basic lack of understanding of human behavior and the perceptual process, it also may reflect a wistful desire of some to reduce response behavior to some simplistic level (or some combination of the two). The facts are that humans are neither invariant nor are their responses to products invariant above and beyond the fact that no two products are the same! Sensory professionals do an injustice to themselves and to the science when they embrace these simplistic notions about human behavior without fully appreciating the consequences. They also do a disservice when they participate as subjects, thereby perpetuating the notion of the expert, the *N* of 1 who can make these absolute judgments.

With the growth of the economy and competition and the evolution of processed and formulated foods, experts faced increasing difficulty in maintaining an awareness of all developments concerning their own product interests. As a further complication, product lines expanded to the extent that it was and is impossible for an expert to have detailed knowledge about all products, let alone the impact of different technologies. While the expert was required to continue making finite decisions about product quality, consumer attitudes were changing in ways that were not fully appreciated. With the development of contemporary measurement techniques and their application to sensory evaluation, it became evident that reliance on a few experts was questionable. To deal with this problem, some companies turned to sensory evaluation (which was often referred to as "organoleptic analysis" in the early literature). In truth, companies did not turn directly to sensory evaluation as a solution to the failure of experts, rather the marketplace created opportunities. As competition increased, and became more national (and eventually, international) in scope, the need for more extensive product information became evident. Managers were either disappointed with results from some types of consumer tests and/or costs became increasingly difficult to justify to management, and now were more willing to consider alternative sources of product information. For those companies

where there were sensory resources, opportunities developed and in some instances considerable success was achieved. To that extent, sensory evaluation represented a new, and as yet, untried resource. Before discussing this contemporary view, it is necessary to further explore the earlier developments of sensory evaluation.

As noted previously, sensory evaluation was of considerable interest in the late 1940s and on into the 1950s, prompted in part by the government's effort to provide more acceptable food for the military (Peryam *et al.*, 1954), as well as by developments in the private sector. For example, the Arthur D. Little Company introduced the Flavor Profile Method (Caul, 1957), a qualitative form of descriptive analysis that minimized dependence on the technical expert. While the concept of a technical expert was and continues to be of concern, the Flavor Profile procedure replaced the individual with a group of about six experts (that they trained) responsible for yielding a consensus decision. This approach provoked controversy among experimental psychologists who were concerned with the concept of a group decision and the potential influence of an individual (in the group) on this consensus decision (Jones, 1958). Nonetheless, at that time, the method provided a focal point for sensory evaluation, creating new interest in the discipline, which stimulated more research and development into all aspects of the sensory process. This topic is covered in more detail in the discussion on descriptive methods in Chapter 6.

By the mid-1950s, the University of California at Davis was offering a series of courses on sensory evaluation, providing one of the few academic sources for training of sensory evaluation professionals. It should be mentioned that other universities, including Oregon State University, University of Massachusetts, and Rutgers offered course work in sensory evaluation but not to the extent as offered by the University of California (subsequently, many other Universities initiated independent courses in sensory evaluation producing more professionals in the discipline). These developments are reflected in the food science literature of this same period, which includes many interesting studies on sensory evaluation by Boggs and Hansen (1949), Harper (1950), Giradot *et al.* (1952), Baker *et al.* (1954), Pangborn (1964), and Anonymous (1968). These studies stimulated and facilitated the use of sensory evaluation in the industrial environment. The early research was especially thorough in its development and evaluation of specific test methods. Discrimination test procedures were evaluated by Boggs and Hansen (1949), Giradot *et al.* (1952), and Peryam *et al.* (1954). In addition to discrimination testing, other measurement techniques also were used as a means for assessing product acceptance. While scoring procedures were used as early as the 1940s (see Baten, 1946), primary emphasis was given to use of various paired procedures for assessing product differences and preferences. Rank-order procedures and hedonic scales became more common in the mid- to late 1950s. During this time period, various technical and scientific societies such as Committee E-18 of the American Society for Testing and Materials, the Food and Agriculture Section of the American Chemical Society, the European Chemoreception Organization, and the Sensory Evaluation Division of the Institute of Food Technologists organized activities focusing on sensory evaluation and the measurement of flavor. For a review of the activities of ASTM, the reader is referred to Peryam (1991).

III. Development of Sensory Evaluation

It would be difficult to identify any one or two developments that were directly responsible for the emergence of sensory evaluation as a unique discipline and its acceptance (albeit, on a limited basis) in product business decisions. Certainly the international focus on food and agriculture in the mid-1960s and on into the 1970s (it is still with us), the energy crisis, food fabrication and the cost of raw materials (Stone, 1972), competition and the internationalization of the marketplace have, directly or indirectly, created opportunities for sensory evaluation. For example, the search for substitute sweeteners stimulated new interest in the measurement of perceived sweetness (along with time-intensity measures). This, in turn, stimulated development of new measurement techniques (Inglett, 1974) and indirectly stimulated interest in development and use of direct data entry systems as a means for evaluating the sweetness intensity of various ingredients (see Anonymous, 1984; Guinard *et al.*, 1985; Gordin, 1987; Winn, 1988 for more information about this latter topic). Today, this situation has not changed and there remain many unfulfilled opportunities. Whether companies are developing new products, attempting to enter new markets or compete more effectively in existing markets, the need for sensory information remains (see Stone, 2002, *Sensory Evaluation and the Consumer in the 21st Century*). While much more could be written and speculated about these opportunities and their antecedents, it is more important that our attention be focused on how this sensory resource should be structured so it can function more effectively in the future.

After a long and somewhat difficult gestation, sensory evaluation has emerged as a distinct, recognized scientific specialty (Sidel *et al.*, 1975; see also Stone and Sidel, 1995). While the focus of this article was on the use of sensory evaluation in the development of fabricated foods, there were implications for sensory evaluation, in general. As a unique source of product information it had important marketing consequences, providing direct, actionable information quickly and at low cost. It was proposed that organizing of sensory evaluation test services along well-defined lines (e.g. formal test requests, selection of a test method based on an objective, and selection of subjects based on sensory skill) would increase the likelihood of such services being accepted as an integral part of the research and development process (or other business units within a company). It has become clearer that without an organized approach, a management-approved plan, and an operational strategy, sensory resources are rarely used effectively and are less likely to have a significant, long-term impact. More recently, some professionals have structured their sensory resources around a single method in the mistaken assumption that it will solve all problems, or that it obviates the need for a more broadly based program with specified goals and objectives and operational plan. As a general rule, reliance on a single method will create more problems than it will ever solve and hinder development of sensory evaluation as an integral part of the business decision-making process. Reliance on a single method is particularly risky because it leads one to modify problems to fit the method, and to overlook basic sensory procedures and practices. In a short course given several decades ago, Pangborn (1979) called attention to misadventures that have occurred in sensory evaluation which included this reliance on a single method. The article was

one of several by this author as part of her continuing efforts to improve the quality of the sensory evaluation literature, now that its use has become more common. The three issues of particular concern to Pangborn were the lack of test objective, adherence to a test method regardless of application, and improper subject selection procedures. These three issues remain as such, even now. These are not the sole property of the sensory literature (and by default many in teaching roles) but also are quite commonplace in the business environment. It is clear that much more needs to be done to improve the quality of sensory information.

An interesting development for sensory evaluation has been the continued growth in the number of short courses and workshops being offered. When there were few University offerings, such programs served a useful purpose for individuals with responsibility for their company's sensory program. In the past decade, there has been a quantum increase in the number courses being offered, including distance learning, which suggests that University offerings are still insufficient for industry's needs. Our own experience in offering courses during the past three decades reflects a continued interest in sensory evaluation, especially the more pragmatic issues of developing a program within a technical-business environment. Some of the material presented in this book evolved from workshop material that has proven especially beneficial to participants. Newspapers and other public information sources present articles about sensory evaluation (not regularly, but often enough to be noticeable). These articles usually include some impressive revelations (to attract the reader) about the special "tongue" or "nose" of certain individuals who are claimed to have almost mystical powers. These individuals are generally associated with such products as wine, beer, coffee, fragrance, or function as wine and food writers. Still other approaches convey an impression that the sensory secrets of the subconscious mind are being tapped by some new technique, with the end result being the ideal consumer product. While sensory professionals should welcome publicity, such promotional information is not always scientifically based (in some instances it may have no basis at all). In spite of this lack of consistency, continued press coverage about sensory evaluation is helpful to the field, if only because it reaches key people who might not otherwise read about it. These changes, and a greater awareness of sensory evaluation, appear to have coincided with a dramatic shift by the food industry toward a consumer-oriented environment and away from the more traditional manufacturing/production-oriented environment. By that we mean a recognition that understanding consumer attitudes and behavior is essential information and ought to be known before one formulates a product rather than manufacturing a product and looking to others; for example, marketing, to convince the consumer to purchase that product.

Opportunities for sensory evaluation continue to develop primarily as a result of significant changes in the marketplace and to a much greater extent than changes in sensory evaluation. Mergers, leveraged buy-outs, and other financial restructuring activities, and the internationalization of the marketplace have created even greater challenges in the consumer products industry. There are numerous choices in terms of brands, flavor alternatives, convenience, pricing, new products, and combinations not thought of a decade ago; for example, yogurt beverages, etc. Many companies have determined that new products at competitive prices are essential for long-term growth

and success. However, this has presented its own unique challenges and risks (see, e.g. Meyer, 1984). New product development and the proliferation of choices within a product category rapidly accelerated in the 1980s at a rate neither appreciated nor believed possible in the past. This acceleration was accompanied by considerable financial risk (Anonymous, 1989). In a publication on the topic, Carlson (1977) determined that the failure rate of new products has, at times, been as high as 98% for all new products. Since that report one can easily find similar reports in the trade literature; clearly, this situation has not changed very much and certainly not for the better. From a business perspective, this risk severely challenges creative skills and available technical resources and has provided renewed interest in other information resources such as sensory evaluation (see Stone, 2002). Companies are now more receptive to new approaches and to new ways for anticipating and measuring the potential for a product's success in the marketplace. Of course, some companies may choose to not introduce new products and thereby minimize much of that risk but to rely on brand and line extensions (Lieb, 1989). Here too, the need for sensory information is essential if such products are to be brought into the market within reasonable time and budgetary considerations. These changes should have accelerated the acceptance of sensory evaluation; however, this has not occurred to any great extent until more recently. Companies are now more aware of sensory evaluation; however, the organization and operation of sensory business units with full management support still lag other related activities such as consumer insights (a successor to marketing research?). Nonetheless, the fact that some programs are fully operational bodes well for the future.

While much progress has been made, considerably more remains to be achieved, particularly within the business environment. In the next chapter, this issue is more fully explored, with particular emphasis on structural (e.g. organizational) issues and their integration with the functioning of sensory resources; that is, how methods and subjects are developed and used to solve specific problems and maximize sensory's benefits to a company.

IV. Defining Sensory Evaluation

To more fully appreciate the situation, it is helpful if we first consider the manner in which the two words *sensory evaluation* are defined. The paucity of definitions is surprising: a perusal of various texts and technical sources reveals one prepared by the Sensory Evaluation Division of the Institute of Food Technologists (Anonymous, 1975) and quoted here, that provides insight into the subject:

Sensory evaluation is a scientific discipline used to evoke, measure, analyze and interpret reactions to those characteristics of foods and materials as they are perceived by the senses of sight, smell, taste, touch and hearing.

This definition represents an obvious attempt to be as inclusive as is possible within the framework of food evaluation, with the word "food" considered global; that is, an

14 Introduction to Sensory Evaluation

ingredient is a food. Similarly, materials can be products for the home such as furniture polish, a product for personal care such as a shampoo, a hair colorant, or a lipstick, etc. While it may be relatively easy to find fault with the narrowness of the definition, we should perhaps look beyond the terminology and consider it in a much broader context. First, the definition makes clear that sensory evaluation encompasses all the senses. This is a particularly important issue and one that is overlooked, with the result that in some environments sensory evaluation is viewed solely as “taste testing,” as if to imply that it excludes the other senses. While an individual may be asked to respond to a particular product attribute, for example, its color, but if no special care has been taken to exclude the product’s aroma, then it is very likely that the obtained color response will be affected by the aroma but not in a predictable way. This will lead to a confounding of the response and potential misinterpretation of the results. A product’s appearance will impact an individual’s response to that product’s taste, etc. Regardless of what one may like to believe or has been told, responses to a product are the result of interactions of various sensory messages, independent of the source. To avoid obtaining incomplete product information, it is important to design studies that take this knowledge into account. The familiar request to “field a test but tell the subjects to ignore the color as that will be corrected later” is a sure sign of impending disaster. This issue will be discussed in a subsequent chapter but is mentioned here to emphasize its importance in the overall evaluation process and the seemingly lack of appreciation of its consequences for the sensory professional reporting results to management who does not “ignore the color.” Second, the definition seeks to make clear that sensory evaluation is derived from several different disciplines, but emphasizes the behavioral basis of perception. This involvement of different disciplines may help to explain the difficulty entailed in delineating the functions of sensory evaluation within the business and academic environments. These disciplines include experimental, social, and physiological psychology, statistics, home economics, and in the case of foods, a working knowledge of food science and technology.

As the definition implies, sensory evaluation involves the measurement and evaluation of the sensory properties of foods and other materials. Sensory evaluation also involves the analysis and the interpretation of the responses by the sensory professional; that is, that individual who provides the connection between the internal world of technology and product development and the external world of the marketplace, within the constraints of a product marketing brief. This connection is essential such that the processing and development specialists can anticipate the impact of product changes in the marketplace. Similarly, the marketing and brand specialists must be confident that the sensory properties are consistent with the intended target and with the communication delivered to that market through advertising. They also must be confident that there are no sensory deficiencies that lead to a market failure. Linking of sensory testing with other business functions is essential just as it is essential for the sensory professional to understand the marketing strategy. In recent years, other business units have expressed interest in using sensory information. For example, quality control/quality assurance professionals have initiated efforts to include sensory information into the quality equation. Here too, it has been found

that sensory information is cost-effective. However, it has its own set of challenges that will be discussed in more detail later in this book. Thus, sensory evaluation should be viewed in much broader terms. Its contributions far exceed the questions of which flavor is best or whether ingredient A can be replaced with ingredient B. In this discussion and in those that follow, it should be obvious to the reader that we assign more responsibilities to sensory evaluation than just the collecting of responses from selected individuals in a test situation.

This concept is especially important in balancing the impact of consumer response behavior as developed by marketing research. In recent years, there has been a growing interest in and much greater use of sophisticated physiological and psychological approaches in measuring consumer behavior. Although the initial interest centered around advertising, it was only natural that it should also be applied to consumers' reactions to the products themselves. While such approaches can be expected to have a positive impact on the product's success in the marketplace in the long run, it would be naive to assume that such information can substitute for all sensory testing. There are those who advocate by-passing sensory testing itself, and relying on larger-scale consumer testing into which a sensory component is incorporated. Such an approach has much appeal because it could save time (and some cost) and works directly with the real consumers. This approach has considerable surface appeal; however, it comes with risk. In any population, there exists a wide range of product sensitivities. In fact, about 30% of any consumer population cannot satisfactorily discriminate amongst products they regularly consume. As we will discuss later in this book, it takes as much as 10 or more hours to teach consumers how to use their senses and as much as 10 more hours for them to learn how to verbalize what they perceive. So bypassing the sensory test by asking consumers to respond to sensory questions or having them verbalize their experiences has considerable risk associated with it. Along these same lines, the tendency to have a universal test method also presents similar challenges to the sensory professional. While it would be extremely useful to have one method to provide all needed information, it is unrealistic to assume such a solution to account for the complexities of consumer behavior. In addition, the trend toward relying on a statistical solution is equally risky when the basic information was obtained without following accepted procedures and practices.

There is no question that there are important and necessary links between sensory evaluation (and technology) and marketing research. However, one cannot substitute for the other, nor should one be designed to include the other without appreciating the risks. The contexts for both test types are different. Sensory tests focus on product formulation differences and their magnitudes, product preferences, and the relationships among the test variables. Sensory also is capable of linking these variables with consumer attitudes, purchase intent, benefits and uses, etc. While marketing research tests will obtain responses to product attributes, these responses are not a substitute for results from a trained descriptive panel. This is not to suggest they have no value, rather they are indicative of a problem when results are not consistent with existing data. This and related issues will be discussed in more detail in Chapter 7. Similarly, the results of a sensory acceptance test are not a substitute for a marketing research test with a larger population of qualified consumers. A failure to appreciate these

differences has had and will continue to have deleterious effects on the growth and development of sensory resources in a company and also on the ability of a company to develop and market successful products. These issues will be discussed again in subsequent chapters of this book.

V. A Physiological and Psychological Perspective

Sensory evaluation principles have their origin in physiology and psychology. Information derived from experiments with the senses has provided a greater appreciation for their properties, and this greater appreciation, in turn, has had a major influence on test procedures and on the measurement of human responses to stimuli. Although sources of information on sensory evaluation have improved in recent years, much information on the physiology of the senses and the behavioral aspects of the perceptual process has been available for considerably longer (Morgan and Stellar, 1950; Guilford, 1954; Granit, 1955; Geldard, 1972; Harper, 1972; Tamar, 1972; Poynder, 1974; McBride and MacFie, 1990). In this discussion, we will identify our meaning of the word “senses.” As Geldard (1972) has pointed out, classically the “five special senses” are vision, audition, taste, smell, and touch. The latter designation includes the senses of temperature, pain, pressure, and so forth. Numerous efforts have been made to reclassify the senses beyond the original five, but they are the ones we have chosen to consider here.

From study of the physiology and anatomy of the systems, we know that each sense modality has its own unique receptors and neural pathways to higher and more complex structures in the brain (Morgan and Stellar, 1950; Pfaffman *et al.*, 1954; Granit, 1955; Tamar, 1972). At the periphery, receptors for a specific sense (e.g. visual and gustatory) respond to a specific type of stimulation that is unique to that system. That is, a gustatory stimulus does not stimulate visual receptors. However, when the information is transmitted to high centers in the brain, considerable integration occurs. Comprehension of how sensory information is processed and integrated is important in understanding the evaluation process (Stone and Pangborn, 1968; McBride and MacFie, 1990). What this means when translated into the practical business of product evaluation is that products are a complex source of stimulation and that stimulation will not be exclusive to a single sense, such as vision or taste. Failure to appreciate the consequences of this very fundamental component of sensory evaluation continues to have serious consequences (a perusal of the current sensory literature provides ample evidence of these practices). Consider an evaluation of a strawberry jam that has visual, aroma, taste, and textural properties. Requiring subjects to respond only to textural attributes (and ignore all other stimuli) will lead to partial or misinformation about the products, at best. Assuming subjects (or anyone for that matter) are capable of mentally blocking stimuli or can be trained to respond in this way, is wishful thinking. Response to all other stimuli will be embedded in the textural responses which, in turn, leads to

increased variability and decreased sensitivity. This approach ignores basic sensory processes and the manner in which the brain integrates incoming information and, combined with memory, produces a response. Probably more harm is done to the science of sensory evaluation and its credibility when procedures and practices are modified based on faulty assumptions about human behavior in an effort to eliminate a problem which usually has nothing to do with behavior. Use of blindfolds and rooms with colored lights are examples that quickly come to mind. Both practices reflect a poor understanding of the perceptual process, and the mistaken belief that by not measuring some product sensory characteristics, they can be ignored! These practices are never recommended, as they are sources of variability that lead to increased risk in decision-making. In subsequent chapters, solutions to these problems are proposed along with some practical examples.

The psychophysical roots for sensory evaluation can be traced to the work of Weber (cited in Boring, 1950), which was initiated during the middle of the nineteenth century. However, it could be argued that it was Fechner (Boring, 1950), building on the experimental observations of Weber, who believed he saw in these observations a means of linking the physical and psychological worlds that gave rise to the field of psychophysics. Much has been written about this topic as well as the early history of psychology in general. The interested reader will find the book by Boring (1950) especially useful in describing this early history. For purposes of this discussion, we call attention to just a few of these developments because of their relevance to sensory evaluation. This is not intended to be a review as such reviews are readily available in the psychophysical literature, rather to remind the reader, in an abbreviated manner, the strong ties between sensory evaluation and experimental psychology. For a more contemporary discussion about the perceptual process, the reader is referred to Laming (1986), McBride and MacFie (1990), and Lawless and Heymann (1999).

Fechner was most interested in the philosophical issues associated with the measurement of sensation and its relation to a stimulus. He proposed that since sensation could not be measured directly, it was necessary to measure sensitivity by means of differential changes. This conclusion was based on the experimental observations of Weber. By determining the detectable amount of difference between two stimuli (the just-noticeable-difference or JND), Fechner sought to establish a unit measure of sensation. He proposed that each JND would be equivalent to a unit of sensation and that the JNDs would be equal. From this point, an equation was formulated relating response to stimulus:

$$S = k \log R$$

where S is the magnitude of sensation, k a constant, and R the magnitude of the stimulus.

As Boring emphasized, Fechner referred to this as Weber's Law, now known as the Weber-Fechner Law or the Psychophysical Law. This initiated not only the field of psychophysics but also a long series of arguments as to the true relationship between stimulus and response and the development of a unified theory of perception. For many years, it was argued that one could not measure sensory magnitudes and, therefore, such a psychophysical law was meaningless. However, the most concerted attacks on

the Fechnerian approach were made by Stevens (1951) and his co-workers (see Cain and Marks, 1971; Lawless and Heymann, 1999, for more detailed discussions on this topic), who advocated a somewhat different explanation for the relationships of stimulus and response. Stevens proposed that equal stimulus ratios result in equal sensation ratios rather than equal sensation differences as proposed by Fechner. Mathematically, as proposed by Stevens, the Psychophysical Power Law was as follows:

$$R = kS^n$$

and

$$\log R = n \log S + \log k$$

where R is the response, k a constant, S the stimulus concentration, and n the modality-dependent exponent.

The formulation of this law had a renewed and stimulating effect on the field, as manifested by hundreds of publications describing responses to a variety of stimuli including commercial products, and setting off numerous debates as to the numerical values of the power functions for the various sensory modalities. Certainly, the development of signal detection theory has had a major impact on our knowledge of the perceptual process, in general, and on sensory evaluation in particular. But, it too has its opponents who seek a single unified theory of perception. However, as Stevens (1962) observed, such laws generally attract criticism as well as support, and the Power Law was no exception (Anderson, 1970). In fact, the extent of criticism directed at any one theory of perception is no greater or less than that directed at any other theory and to date, there has been no satisfactory resolution. The interested reader should read the previously cited work of Anderson and Laming, and for more on this issue from a sensory evaluation perspective, see Giovanni and Pangborn (1983). The observations of Weber, however, warrant further comment because of their importance in product evaluation. Basically, Weber noted that the perception of the difference between two products was a constant, related to the ratio of the difference, and expressed mathematically as

$$K = \frac{\Delta R}{R}$$

where R is the magnitude of the stimulus and K is the constant for the JND. Experiments on a variety of stimuli and particularly those involving food and food ingredients, have shown generally good agreement with Weber's original observations (Wenzel, 1949; Schutz and Pilgrim, 1957; Luce and Edwards, 1958; Stone, 1963; Cain, 1977). As with other developments, there are exceptions and not all experimental results are in complete agreement with this mathematical expression. Nonetheless, the JND has found widespread application in the sensory evaluation of products as described in Chapter 5.

Among the contributions to sensory evaluation by psychologists, the work of Thurstone (1959) is especially noteworthy. Many of his experiments involved foods (rather than less complex unidimensional stimuli) or concerned the measurement of attitudes, both topics that are of particular interest to the manufacturer of consumer

goods. Thurstone formulated the Law of Comparative Judgment enabling use of multiple paired comparisons to yield numerical estimates of preferences for different products. From this and other pioneering research evolved many of the procedures and practices used today by sensory professionals. Today, renewed interest in Thurstonian psychophysics is welcomed; however, these authors never considered it absent or not an integral part of the perceptual process.

Psychology has contributed substantially to our understanding of the product evaluation process; however, it would be incorrect to characterize sensory evaluation solely as a part of the science of psychology.

As the adherents to the various schools of psychology or particular approaches espouse their causes, it is easy to confuse research on scaling or perception with assessment of product success without appreciating the differences and the risks. It is easy to mistake developments in psychology as having immediate and direct consequences for sensory evaluation, and to blur the lines between the two. This has been particularly evident by those who implied use of magnitude estimation, a type of ratio scale, as the best procedure for obtaining meaningful numerical estimates of the perceived intensities of various stimuli (Stevens, 1960). It was claimed that use of this scaling method enabled one to obtain unbiased judgments, allowed for use of higher-order statistics, and a much greater level of sensitivity than achievable by other scales. As will be discussed in Chapter 3, attempts to demonstrate this scale superiority in sensory evaluation applications have been unsuccessful (Moskowitz and Sidel, 1971). What may be demonstrated with a single stimulus in a controlled experiment does not necessarily translate to the evaluation of more complex stimuli such as foods and other consumer products. That such superiority has not yet been demonstrated with any scale should not be surprising, particularly from an applications perspective. As Guilford (1954), Nunnally (1978), and others have noted, having a less than perfect scale does not invalidate results and any risks that such use entailed are not as serious as one might expect. As it happens, the risks are very small, particularly with regard to the product evaluation process. In Chapter 3, this topic will be discussed in more detail. Nonetheless, it should be clear that sensory evaluation makes use of psychology, just as it makes use of physiology, mathematics, and statistics in order to achieve specific test objectives. The discipline operates without the restrictions of different theories of perception competing to explain the mechanisms of perception, a task of psychophysicists and physiological psychologists. The sensory professional must be sufficiently well-informed in these disciplines to identify their relevance to the organization, fielding, and analyses of a sensory test, without losing sight of sensory goals and objectives.

The remaining chapters of this book emphasize the organization and operation of a sensory evaluation program, taking into account both the short- and long-term issues that directly affect the success of a program.

