



The Origin and Evolution of Human-Centered Food Product Research

1

Howard R. Moskowitz and Herbert L. Meiselman

Contents

Historical Background: The Early Twentieth Century	4
A Century Ago or Even Earlier: Before Consumers	4
The Early Days, Product Quality, Statisticians, and Statistical Tests	5
Quality Control and the Rise of Subjective Testing	5
Statisticians, Difference Tests, and the Emergence of Interest in Design of Experiments (DOE)	7
Describing Versus Designing: Structuralist Versus Functionalist Agendas	8
By Way of Introduction to this Section	8
Contrasting Beliefs and a Detour: What Can Be Judged and by Whom	9
A Land of Plenty: The Rise of Descriptive Analysis (Structuralism) to Guide Product Design	10
Psychophysics (Functionalism) Moves into the Food Industry	11
A Cadre of Chemosensory Psychophysicists Enters the World of Food	13
The Zeitgeists of “Disciplines”: Psychophysics, Sensory Professionals, and Market Researchers	14
Zeitgeists of Method in the 1960s, 1970s, and 1980s	15
Single Test Stimulus and Analysis: Cross Tabulations	16
Single Test Stimulus: Just About Right (JAR) Scales	16
Single or Multiple Test Stimuli: The Self-Designed Ideal	17
The JAR Scale, the Self-Designed Ideal, and Efforts to Validate Them	17
Multiple Test Stimuli: Mapping	19
Multiple Test Stimuli: Response-Response Analysis	20
Systematics: Creating and Using Psychophysical Curves	21
Stimulus-Response Analysis	21

H. R. Moskowitz
Mind Genomics Associates, White Plains, NY, USA
e-mail: mjihrm@gmail.com; mjihrm@sprynet.com

H. L. Meiselman (✉)
Herb Meiselman Training and Consulting, Rockport, MA, USA
e-mail: herb@herbmeiselman.com

Expectations	23
Conclusions	24
References	26

Abstract

This chapter presents a history of human-centered product research including product design and product evaluation. The chapter begins with the world of food business in the early 1900s and moves through a variety of themes and disciplines, including the period of testing for quality, the reign of the “golden palate,” the emergence of descriptive analysis, and the growth of psychophysical thinking. The chapter traces the history of this field up until the growth of modern methods beginning in the 1970s. The chapter shows the approaches, some of the early research pioneers, some of the types of results obtained, an analysis of “whys” behind the themes, and the controversies based on intellectual histories.

Historical Background: The Early Twentieth Century

If one were to look at the way products are developed today, one might think that statistics, modeling, professional expertise, and a strong intellectual heritage have been in play for decades. Most food companies, worldwide, subscribe to the now widely held belief that to design foods requires a knowledge of the consumer, a talent in the creation of tasty mixtures, and a profound understanding of trends, all together collaborating to make today’s test products become tomorrow’s massive successes. One cannot help but be impressed by the talents involved, whether one refer to expert panels (e.g., Caul 1957), to statistical texts (Box et al. 2005), to consultants specializing in new product design, or to trends monitoring which drive the development of new foods (Costa and Jongen 2006).

This review covers the various *Zeitgeists* (German term literally meaning “spirit of the time”), in the history of human-centered research, including both product design and product evaluation, beginning with the “golden palate,” and moving to today, with the plethora of product offerings, the paradox of choice, the change in the nature of expertise, and recently the focus on low cost testing.

A Century Ago or Even Earlier: Before Consumers

If one wants to understand how today is so different, then one can go back a century or more, to a time when the packaged goods industry for foods was just developing (► Chap. 57, “Measuring Meaning of Food in Life” by Ruffieux in this volume also discusses this early period). The time was decades just before the turn of the century to some 10 or 15 years after World War I. Those with a penchant for photographic history are treated to the myriad photographs of the stores those days, the corner

grocers, often weighing out the product, in a way that can be still seen when one goes to some stores or to farmer's markets. The shopper could buy some packaged goods, but there were few supermarkets, and even fewer, if any, elaborate food halls in the way we now know. Of course, big cities had big food markets even years ago, to entice shoppers and make the effort easy and social. The storekeeper might offer the customer a taste of the product before purchase, a delightful custom which remains today, but today's in-store tastings are now efficient affairs, with companies specializing in in-store demonstrations and tastings.

Nonetheless, in such a world product development proceeded with commercial products like Kellogg's Corn Flakes (started 1906) and Campbell Soup Company (started 1869, condensed soups 1896–1897), along with other iconic companies and brands. The world of iconic brands stretched all over Europe, from Twinings and Cadbury in the UK to Liebig in Germany and so forth. The focus was on the one product to sell, not on the paradox of choice driven by a plethora of products which occupy the shelf in a marketing war, so-called facings, each item being a different size of the basic product, a different flavor, and so forth, all designed to proclaim "Buy Me," and grab the customer's last money.

The Early Days, Product Quality, Statisticians, and Statistical Tests

Quality Control and the Rise of Subjective Testing

In the 1930s and 1940s, product design was just a gleam in the mind of some forward thinkers. The notion that one could formalize product design using subjective perceptions was new and only infrequently recognized, much less implemented. Typically, practitioners used sensory perception to determine whether two products were the same or different (e.g., two batches from the product, two samples from different suppliers). This seemingly simple task laid the groundwork for product design, not so much because difference tests guide design, but because they paved the way to use subjective measures to guide decisions about products, such as accept/reject. What began as new efforts by some bench researchers and developers to understand differences would lay the groundwork for today's product design efforts. The full flower, however, would take years, as well as the entry of many different talents, such as marketers as well as experimental psychologists, especially psychophysicists.

The individuals who were doing tests of acceptance in the late 1930s and early 1940s, worked in companies and, later, in the late 1940s, were at the Food and Container Institute of the Armed Forces operated by the US Army Quartermaster Corps in Chicago. From these professionals and others who attended the yearly conferences of the IFT (Institute of Food Technologists), one got a "picture" of sensory evaluation in the guidance of product design.

Among the biggest contributors to the development of the field of human-centered research on food products were the US Army and its laboratories, first in Chicago, Illinois, and then in Natick, Massachusetts. The Army laboratories in

Chicago and Natick are extensively reviewed by Meiselman and Schutz (2003). The Army's Quartermaster Subsistence Research and Development Laboratory in Chicago established a Food Acceptance Branch in 1944, headed by Franklin Dove. This was followed by a food acceptance conference in December 1945, probably the first food acceptance conference worldwide. Dove published a paper on food acceptability in 1947 which describes the new taste panel booths at the laboratory. In 1949, David Peryam became head of the Food Acceptance Branch. Peryam was a psychologist, marking the introduction of psychologists into the field of food acceptability. Peryam was joined by psychologists Norman Girardot, Howard Schutz, Joseph Kamen, and others. A long line of research psychologists followed Peryam leading up to the current authors of this chapter who joined the Army's laboratory in Natick in 1969.

The work of the Food Acceptance Branch in Chicago focused broadly on food-related behavior, including both acceptance testing of products and studies of food habits. For acceptance testing, the laboratory used both expert and trained panels and consumer panels, usually using a paired preference method (Dove 1947). Studies of food habits were conducted under contract by universities who studied food preferences for a list of foods, along with consumer information on preparing and serving these foods. The studies of food preference continued from the Chicago period to the Natick period.

The most well-known output of the Food Acceptance Branch in Chicago was the development of the nine-point hedonic scale. This project was started by Peryam and Girardot in 1949 (Peryam and Girardot 1952) who were joined by researchers at the University of Chicago in 1951, including L. L. Thurstone, Lyle Jones, and Darrell Bock (Jones et al. 1955). The team conducted many individual studies including testing the semantic meaning of various phrases ("like extremely," etc.), varying the number of categories, scale balance, numbers of positive and negative scale categories, and presence of a neutral category. They found an advantage for longer scales, but no advantage for a neutral point or the same number of positive and negative categories. The nine-point hedonic scale has probably been the most widely used scale of food acceptance in the world, because it is easy for consumers to use. Many people using it do not know the extensive testing that went into its development. Nevertheless, the scale has had problems because of problems with translation into other languages (ref) and because of the lack of equal intervals as should be found on an interval scale (see ► Chap. 12, "Measuring Liking for Food and Drink" by Ares and Vidal in this volume) (Fig. 1).

Both Chicago and Natick Laboratories worked on the relationship of food acceptance to food consumption. Schutz and Kamen found that 50% of the variance in choice or consumption could be accounted for by hedonic mean scores but that up to 75% could be accounted for if one looked at individual food groups rather than food items. Cees de Graaf and colleagues (de Graaf et al. 2005) found significant but moderate correlations using data from Natick field testing of rations,

Like Extremely
Like Very Much
Like Moderately
Like slightly
Neither Like nor Dislike
Dislike Slightly
Dislike Moderately
Dislike Very Much
Dislike Extremely

Nine point hedonic scale: in its pure form the 9 point hedonic scale does not have lines connecting the scale points and does not have numbers associated with each scale point.

Fig. 1 Nine-point hedonic scale

indicating that acceptance is an important factor, but not the only factor contributing to intake.

Statisticians, Difference Tests, and the Emergence of Interest in Design of Experiments (DOE)

The emergence of interest in testing differences naturally involved statisticians, who were the experts regarding inferential statistics. By the period between the late 1930s and the 1950s, the world of food design and development was becoming more professionalized. There was interest in doing the “right test.” For the first few years, the focus was on inferential statistics, namely, do two or more products differ from each other? The question of the time was “is there a difference, and is the difference making the product better, or is it making the product worse?” Product designers and developers were not thinking about systematic variation of products, nor about uncovering rules. The times called for the human being simply as a tester, an “assessor,” in many of the scientific papers.

Statisticians, however, were moving to modeling, in order to understand the relation between variables. A new culture was beginning to emerge, a culture of systematic variation. Statisticians realized that they could use their armory of tools to understand the relations between variables, specifically those variables under the developer’s control, and other variables measured by people, namely, perception. For example, one could learn how the amounts of two ingredients, e.g., sweetener and flavoring, together drive the perception of “perceived strength of flavor” by systematically mixing different levels of sweetener with different levels of flavoring (the ingoing experimental combinations or experimental design) and instructing panelists to rate the flavor intensity as they perceived it. This systematic approach produced a body of useful knowledge for the developer, far deeper than the knowledge that would be obtained by asking why two beverages of the same type differ in their respective “strengths of perceived flavor.”

The foregoing, focusing on systematic variation, and the linkage between two variables, an independent variable under control and a dependent variable that is measured, falls into the category of designed experiments, or more typically and formally DOE, Design of Experiments. The application of DOE to the design of food would come later, in the late 1940s, more in the 1950s and 1960s, and flower in the 1970s–1990s. During 1950–1965, there was a growing interest in DOE in the chemical industry, perhaps because it had financial ramifications such as increasing the yield from chemical reactions.

DOE opened the minds of professionals, allowing for more complicated experiments, and thus laid the groundwork for product design. The traditional scientific method postulated how important it was to isolate a variable and study it. The notion that one should look at mixtures, to understand the system, was occasionally thought of, but the lack of fast and easy computation made these multivariate studies more of theory to be studied than a work tool to be used. DOE would become far more important later, when computers enter the science, because DOE PLUS COMPUTERS enabled the statistician, and thus the engineer, and scientist, to study the workings of several variables at once, and even nonlinearities and interactions between variables. All these would become important in the world of food design. Names such as Cochran (1950) and Box et al. (2005) are important, as are Plackett and Burman screening designs (Vanaja and Shobha Rani 2007). These statisticians contributed to this newly emerging field of DOE, which would lead to food design some decades later. The reader is directed toward the later papers cited here, because they put the earlier work into clear focus in terms of the major contributions made in those early methods.

Describing Versus Designing: Structuralist Versus Functionalist Agendas

By Way of Introduction to this Section

Structuralism and functionalism are two early schools of experimental psychology, holding sway from the end of the nineteenth century to the early and middle part of the twentieth century (Benjamin 1988). Their use here is metaphoric. Structuralism asserted that one could “understand” the psychology of perception by understanding the attributes or aspects that people perceived. To researchers in structuralist psychology, the prescribed approach was introspection, to list the different attributes of perception, whether those be the attributes or dimensions of vision, hearing, smelling, tasting, and feeling. One could get a sense of how people organized their perceptions. In contrast, researchers in functionalist psychology asserted that a better approach was to understand how the person behaved and how different aspects or dimensions of perceptions “functioned” to guide behavior.

Structuralism seems always to precede functionalism, going back to Aristotle. Aristotle's science classified to understand how aspects of the natural world distributed themselves. Yet even for a genius such as Aristotle, knowing the aspects of the world, the different features shared or not shared by living organisms, did not tell Aristotle how these organisms functioned. One could only guess about function by knowing structure.

The same distinction between structuralism and functionalism applies to the design of foods, with the structuralist agenda of description virtually always preceding the functionalist agenda of determining relations between variables. As we see below, the world of food design was dominated in the early days by description, holding the belief that if one "knew" the different notes or perceptual characteristics of the product, one might be able to develop better products of its type, or perhaps correct some of the quality errors. In contrast, later scientists and practitioners developed the relations between ingredients/process levels and consumer acceptance, in order to drive better product design.

Contrasting Beliefs and a Detour: What Can Be Judged and by Whom

A sense of what "was" sensory analysis, and what was its focus, comes from the now historic book by Amerine et al. (1965), *Principles of Sensory Evaluation of Food*. The 602-page book was the first book in the set of monographs published by Academic Press. This book, published more than a half century ago in the mid – 1960s, emerged from years of painstaking library research as well as personal experience. The topics in the Amerine et al. book deal less with the design of foods and more to with the study of the assessors. Examples of chapter titles are Factors Influencing Sensory Measurements (Chap. 5), Laboratory Studies: Quantity-Quality Evaluation (Chap. 8), Consumer Studies (Chap. 9), Statistical Procedures (Chap. 10), and Physical and Chemical Tests Related to Sensory Properties of Foods. The Amerine et al. (1965) book helps us understand the background out of which product design emerges. If one were to trace the history of product design back to first-order questions, perhaps the first question would not be about the product itself but rather about the person who is doing the judging. The first question was "who is able to judge the product?" The question about what the product should be hardly emerges in the early days.

The detour in consumer-driven product design comes out of a continuing issue in the evaluation of food, namely, who is competent to judge the aspects of food and thus give direction to design. Some of this focus on the "expert" can be traced to the world of certain kinds of products, such as wine, beer, and perfume. These products came with a mystique, the wine expert, the *beer meister*, and the expert perfumer. There was a sense that only these experts "knew" the product. The patronizing undercurrent was that for the most part, consumers simply did not know good from

bad, although of course they knew what they liked and disliked. As of this writing, 2019, the role of “experts” continues to be important, especially in the aforementioned areas of wine, beer, and perfume, where sensory properties can be romanticized in advertising, in turn increasing the value of the brand.

A Land of Plenty: The Rise of Descriptive Analysis (Structuralism) to Guide Product Design

Product design would “somewhat” change in the 1930s, as the world enjoyed the bounty of better food, through advances in food preservation and food transportation. An important step was taken by the Arthur D. Little Company, a technical consulting company in Cambridge, Massachusetts, where Stanley Cairncross and his colleagues developed a system for describing the sensory characteristics of a product (Cairncross and Sjöström 1950). The method was being worked on in the late 1930s and would become a flagship approach in the 1940s. The notion was to help product design by identifying the sensory characteristics, so-called “notes” of a product. It was assumed that the product developer would “know what to do” once the notes were identified in the Flavor Profile. Importantly, the exact linkage between this description and product development could not be specified as a series of specific operations, which converted these “notes” to formulations.

The logic of identifying “notes” or attributes, and assuming that such identification, either by experts or consumers would guide product knowledge and thus product development did not begin with the Arthur D. Little Corporation. As noted above, describing one’s sensation appears to be the first step in systematic science, going back to Aristotle’s classification of animals, plants, and even constitutions of city states. Then there was Francis Bacon and, in psychology, the emergence of the Structuralist School, which assumed that we would know how perception works if we could only describe the perceptions that we have. The efforts did not end there. It was assumed in wine making, beer brewing, and fragrance development that a description of one’s perception would somehow lead the developers to design and develop a better product. Again, as noted above, these descriptions resided in the purview of the expert, whether the business-oriented expert (e.g., perfumer) or the trained expert.

Modern-day efforts for product design using descriptive analysis have focused on training panels, including the Texture Profile (Civille and Szczesniak 1973), the Spectrum™ method (Meilgaard et al. 1999), and Quantitative Descriptive Analysis (QDA®; Stone et al. 1974). There remains little published evidence showing the *precise* steps in the linkage between descriptive analysis and product design. Professionals in the world of sensory analysis have, for the most part, directed their use of expert panels to quality control. The experts, trained in descriptive analysis, can describe two samples, and identify the sensory aspects, the notes which make these samples different from each other. This ability to describe the nature of differences can, of course, be a hint to copy a product by incorporating the

“notes,” but it’s more often used to identify the nature of differences between a “gold standard” product (the “ideal”) and a production or storage sample.

Psychophysics (Functionalism) Moves into the Food Industry

In the early 1860s, the German polymath scientist, Gustav Theodor Fechner, hypothesized that one could measure the perceived intensity of a stimulus, doing so by measuring successive difference thresholds (Stevens 1961). That is, one could begin with a sample of salt water and find the concentration of salt water that would be just noticeably different. This magnitude of change was defined as one JND, one just noticeable difference. One could erect a “sensory scale” by cumulating these JNDs and plotting them against the physical level of a test stimulus (e.g., the salt concentration in the water).

The foregoing is an academic treatment of the foundation of psychophysics. It has little or nothing to do with product design, but successors in psychophysics would have a great deal to do with design, and with the success of products. What Fechner suggested is that one could relate the physical intensity of a stimulus (think “ingredient”) to the perceived intensity.

It would be about 80 years until a more direct approach would take hold, one which would link practical product design to psychophysics. This more direct approach is called *direct scaling*. The respondent is exposed to an array of test stimuli, one stimulus at a time in irregular order, and assigns a number on an attribute scale to match the perceived magnitude of the stimulus. The attribute could be sensory, such as the saltiness of salt solutions, or hedonic, such as the liking of the salt solution, or even a more cognitively complex phrase such as “perceived healthfulness of the salt solution.”

Psychophysics/scaling entered the food industry in a slow but relentless progression. The Army psychologists in Chicago, Peryam, Pilgrim, etc. worked with the scaling expert Thurstone at the University of Chicago to develop the hedonic scale and to test many variations of that scale. At the US Army Chicago laboratory, Pilgrim, Schutz, and Kamen conducted a number of studies in taste and odor psychophysics in the 1950s. They conducted research on difference thresholds for the basic tastes (Schutz and Pilgrim 1957a), as well as on the relative sweetness of a number of natural and synthetic sweeteners using suprathreshold rather than absolute threshold measurements, demonstrating the importance of suprathreshold rather than threshold measures (Schutz and Pilgrim 1957b). The first systematic study of interactions of suprathreshold taste stimuli was conducted, and it was found that in most cases the effects were those of simple enhancement or masking. A one-person olfactorium was constructed, and studies included olfactory adaptation and development of an odor classification system.

When the Chicago laboratory moved to Natick, none of the researchers moved with it, and most of them went into industry and did not publish many research papers (see Meiselman and Schutz 2003). The exception was Howard Schutz. Schutz had received his education in experimental psychology in the 1950s. Schutz

was interested in statistical modeling and analysis, rather than in rigorous testing. Schutz's contribution may be to have opened the eyes of the field to the importance of consumers, perhaps because of his tenure at Hunt Wesson Corporation, where he was responsible for running a commercial sensory evaluation laboratory in the 1960s. Schutz moved to UC Davis after Hunt Wesson and spent many summers working at the Natick Laboratory – he was the only Chicago alumnus in sensory and consumer testing who worked at Natick.

Roland Harper was probably the first psychologist in Europe to work on the sensory properties of foods beginning in the 1940s, around the same time as the US Army Chicago group. Harper joined the agricultural laboratory in Shinfield from 1946 to 1950 (where author Meiselman worked in 1990–1991 along with Howard Schutz). There is evidence of Harper's communication with Thurstone in 1948, around the same time that Thurstone was working with the group of Army psychologists in Chicago (Land 1988). In 1961–1962 he spent a sabbatical year in S. S. Stevens' psychophysical laboratory at Harvard University, from which author Moskowitz graduated 8 years later in 1969.

By the late 1950s, another sensory scientist interested in food emerged in Europe, Egon (Ep) Koster in the Netherlands. He also received his degree in sensory psychology, specializing in the sense of olfaction. Later he would apply his sensory and psychology skills to working on food products. Koster contributed important papers in sensory and consumer psychology of food well into the 2000s. Both Harper and Koster were trained in psychology, while another influential early figure was trained in food science.

These later years were the same years that Rose Marie Pangborn was starting her career in the United States. She attended New Mexico University (degree 1953) and then Iowa State University (degree 1955). She began working at the University of California at Davis in 1955, where she worked for 35 years until her death in 2000. UC Davis became a powerhouse in sensory and consumer science during and after her tenure. Her early papers (maiden name Valdes, married name Pangborn) were published in 1956 and 1957.

Drewnowski (1993) reviewed the contributions of Rose Marie Pangborn at the first Pangborn Sensory Science Symposium in 1992. He noted that her work covered sensory evaluation of food and the evaluation of food preferences. Pangborn was one of the first sensory researchers who moved from model systems (sugar in water, salt in water) to real foods, using canned apricots (Valdes and Roessler 1956) and vanilla ice cream (Pangborn et al. 1957). She also studied more complex sensory stimuli, combining several sensory modalities (sweetness, viscosity, texture). Pangborn also was among the first to study individual differences in perception, applying this to product perception and to the relationship to food preference and management of body weight. This interest in individual differences also extended to determinants of food acceptance, and Pangborn included individual attitudes in her research on nutrition. Perhaps the biggest contribution of Rose Marie Pangborn was her pioneering efforts in training 30 years of undergraduate and graduate students to do testing in a logical, scientific, rigorous way and to report the data in the proper format. Pangborn produced a generation of good researchers and teachers.

A Cadre of Chemosensory Psychophysicists Enters the World of Food

The 1960s grew into a fertile period for the growth of interest of psychophysicists in the chemical senses, taste and smell, and in the lower senses, touch. The focus would first be limited to so-called model systems, a focus that would later evolve to real foods and even full meals. The impetus for this was what has been called the “new psychophysics.” This new psychophysics attempted to uncover quantitative relations between the physical stimuli and subjective responses, with many of the results suggesting that the relation could be described by a power function of the form: sensory rating = k (physical intensity)ⁿ. It was the exponent which was of interest (Stevens 1975).

Natick Laboratories was opened in 1954 as the Quartermaster Research and Development Center, and the US Army food research program moved from Chicago to Natick in 1963. The nutrition research from Chicago moved to Colorado and eventually to Natick, and the dietetic services moved to Virginia. Between 1963 and 1966, a taste test laboratory with 11 testing booths was constructed in Natick. A joint annotated bibliography on acceptance and preference research was published jointly by Chicago and Natick (Bell et al. 1965). Beginning in 1966, Harry Jacobs began a program in behavioral sciences (psychology and other human sciences) with a strong emphasis on food. Harry Jacobs hired Linda Bartoshuk from Carl Pfaffmann’s sensory laboratory at Brown University. Jacobs was interested in basic animal studies of appetite regulation. Bartoshuk was interested in basic human studies of taste processes and taste perception. Shortly after that, Howard Moskowitz and Herbert Meiselman were hired. Both Moskowitz and Meiselman would eventually apply their research to actual foods, with Moskowitz interested in the relation between ingredients and perception and Meiselman interested in meals in real-world settings out of the laboratory. Over time a number of other professionals/psychologists worked at Natick including Richard Bell, Armand Cardello, Barbara Edelman-Lewis, Dianne Engell, Edward Hirsch, F. Matthew Kramer, Owen Maller, and Richard Popper. These professionals were joined by a cadre of technical support people and by a rotating group of Army psychologists assigned to Natick for a two-year period, often bringing with them valuable skills in related disciplines. The military psychologists were especially helpful in conducting large-scale studies of food products and food service with military personnel on military bases and in the field. Finally, the civilian staff and the military staff were joined by a large number of visiting scientists from laboratories all over the world.

The Natick researchers conducted basic psychophysical studies of taste, smell, and texture and their relationship to liking. They applied related methods to study food preferences, food compatibilities, and boredom as applied to menu planning. They extended the study of short-term food preferences to studying long-term preferences. They studied the relationship between data collected in the laboratory and data collected in the field from soldiers and from university students. They extended this work with research on the role of context or environment on ratings of food and beverages. This topic would become a major topic in the field decades later

(see Meiselman 2019). Another new approach in sensory and consumer research was the introduction of expectation theory and methods – this topic is covered below.

In both Europe and America, the psychophysicists were just beginning to look at the relation between the physical stimulus and liking, first with model systems as science (e.g., sugar in water), but insensibly moving toward the evaluation of real foods. Their work was primarily academic, but both authors moved inexorably toward the study of actual foods. For example, fairly early in his tenure at the US Army Natick Laboratories, author Meiselman began to measure liking both in the laboratory and in the field (see Meiselman and Schutz 2003). In contrast, during the same period, author Moskowitz began to use psychophysics to study the nature of liking as well as sensory perception as they are driven by the interactions of sweeteners in cola (Moskowitz et al. 1979). In a parallel path, psychophysics was at the ground in one of the first syntheses in behavioral economics, the economics of sweetness scales (Moskowitz and Wehrly 1972). These latter psychophysics studies would remain academic for the early part of the 1980s but then evolve into larger studies, using psychophysical thinking and experimental design. Those will be covered below.

The Zeitgeists of “Disciplines”: Psychophysics, Sensory Professionals, and Market Researchers

The German philosopher Hegel postulated the ever-repeating dialectic of thesis, antithesis, and synthesis. Advances in every era bring with them conflicts and then synthesis and advance in the wake of those conflicts. We turn now to the 1970s, when forces would interact with each other in ways that might now have been expected from the decades before but which would prove extraordinarily fruitful as one looks back at these conflicts. The focus of the period was “drivers of liking,” or in effect, what makes a product good? We move past the era of difference testing, a worldview which remains with us in full force but generally relegated to quality assurance.

In the 1960s, a new discipline emerged, market (or marketing) research. The focus of this new discipline was the consumer and the market. Most of the efforts focused on the role of consumer responses to product advertising and to product packaging at the point of sale. Market researchers combined scientists and practitioners, with an array of learned journals, such as the *Journal of Market Research*, to memorialize some of the more important efforts of an academic nature. There was also Robert Ferber’s book, *The Handbook of Marketing Research* (Ferber 1974) and the comprehensive *Handbook of Marketing Scales* (Bearden et al. 1993).

Market researchers were interested in people, specifically in the response of people to products. It was market researchers who popularized the notion that product acceptance might vary, some of the variance traceable to error inherent in measuring subjective responses, but perhaps also traceable to the fact that people simply had different preferences. Market researchers were able to recognize these differences in their “product tests,” evaluations of product samples before the

product would be launched, to guard against market failure (so-called disaster checks).

With the introduction of market researchers into the world of product testing, there were three different groups of professionals offering direction and “insight” on the design of products, inputs which were sorely lacking just a decade before:

1. *Sensory psychophysics*: Psychophysicists were entering the business world, choosing to work for manufacturing companies or consulting groups. Often, the psychophysicists would work in the sensory department of a company. Abandoning the traditional academic route, these business-oriented scientists often brought with them the desire to use their psychophysics to drive product creation, rather than to continue the clerical work of difference testing, to which the sensory department had evolved.
2. *Sensory professionals*: The sensory professionals remained bound to their descriptive analysis and graphical representation of the data. They seem not to have been able to show the way that descriptive analysis would drive improved product acceptance. Perhaps their main contribution was to pronounce in their presentations that the product developer would (somehow) recognize certain departures from the standard for current quality control or recognize new notes in products that were to be copied, used as springboards for the company’s new and competitive entry.
3. *Market and consumer research*: The discipline of marketing research was growing. The focus, limited as it was on whether a product passed a specific level for acceptance or was better than a comparison product, necessarily had to look at differences among respondents, the individuals who participated in the market research studies. Whereas psychophysicists and sensory professionals focus at the product itself, marketing researchers focused on the consumer and the resulting pattern of acceptance, recognizing that people differed. They did not, however, have the necessary tools to understand what drives liking beyond the so-called cross-tabulation methods, comprising tables of product scores by subgroups of respondents (e.g., frequent vs infrequent users, older vs younger, brand-loyalists vs non-loyalists, and so forth). As part of the contribution of market and consumer research, there would spring up a new form of understanding consumers, so-called psychographic segmentation (Wells 1975). The focus would be on the psychology of the consumer, specifically what type attitudes were possessed by different segments of consumers and what types of behaviors were exhibited. Food was no longer simply the ingredients and taste, but rather the “right message to the right person” and “the right product to the right person.”

Zeitgeists of Method in the 1960s, 1970s, and 1980s

The three groups of players in the design of product used different tools. We will look at the research tools and then discuss the underlying rationale of each tool, the role of the relevant group, and finally how each tool helped move forward the capabilities of food design.

Single Test Stimulus and Analysis: Cross Tabulations

The term “cross tabulation” refers to the evaluation of one product (or perhaps two), deconstructed into the ratings by different groups or different test conditions. As strange as that might sound today, as of this writing (2019), when cross tabulations were done for product tests, mainly by market researchers, there was a sense that the “answers were in there, in the cross-tabs.” A good analogy today is “Big Data.” There is no reason to assume that the analysis will produce an answer telling the product design what to do to create a better product, but it is satisfying to the researcher to show an “effect,” e.g., that more frequent users prefer the less sweet product.

Single Test Stimulus: Just About Right (JAR) Scales

The JAR or just about right scale is widely used today to identify what to change. The JAR scale asks the respondent to judge whether a product has too little of an attribute (e.g., sweetness), just the right amount, or too much.

Although there have been many developments using the JAR scale, such as penalty analysis (Narayanan et al. 2014), the origin of the JAR scale as used today may have its origins in a discussion about applying psychophysics to business issues in product design. During the September 1968 meeting with Loren B. Sjostrom and Anne Nielsen of Arthur D. Little, Inc., author Moskowitz suggested that the Flavor Profile could be improved by using psychophysical scales to identify how to change a product. The researcher would instruct the respondent to rate the amount of a sensory attribute and the degree of change to make a better product. The psychophysical scale would then show the sensory level and in turn the physical level to optimize acceptance, in the opinion of the respondent. Four years after that first 1968 meeting, the approach was codified in a peer reviewed paper in the *Journal of Applied Psychology* (Moskowitz 1972), demonstrating the practicality of the approach with Kool-Aid, tuna fish spread, and hamburgers of different grinds, respectively.

The JAR scale was and remains attractive. It was easy to apply, to analyze, and to report to product developers who could understand what the scale meant. What was not so clear was what exactly to do with the results when the data fail to be accompanied by a functional relation between “sensory amount” and “physical level,” i.e., when the psychophysical curve was absent. For example, when the respondents said much too sweet, just what did that mean? And what should one do when people disagree? And what about certain attributes for which one never has enough, such as “natural flavor?” In commercial applications for manufacturers conducted as far back as the 1980s, author Moskowitz discovered that for some attributes such as “real chocolate flavor,” the more chocolate one added to the product, the more bitter the product tasted, and the less natural the product tasted. The same type of finding emerged for flavor. Real flavor did not come from the flavoring but from the “sugar.” It required product development expertise to

understand just exactly what the JAR data required in terms of subsequent product design.

Single or Multiple Test Stimuli: The Self-Designed Ideal

Once we admit of the ability of consumers to point to changes in a product, e.g., by the JAR scale, it is not far afield to instruct consumers to describe their ideal product, using the same attributes and the same scale as they used to describe products. One can then compare the magnitude of the ideal product (emerging from the mind of the consumer) to the scores of actual products, tested by these same consumer respondents, at the same time. The products which score closest to the self-designed ideal are presumed to represent target products.

The JAR Scale, the Self-Designed Ideal, and Efforts to Validate Them

The JAR scale and the self-designed ideal instruct consumers to rate products and conceptualize how they would change the product. When it comes to validating the results, how does one then validate the JAR scale and the self-designed ideals, respectively, either analytically or in subsequent direct tests.

Some published literature, primarily in the academic world rather than in the corporate world, has focused on the ability of the JAR scale and the self-designed ideal to guide successful product design. We thus rely on the literature, which presents studies that can be described primarily as methodological. The published studies suggest that the JAR scale and the self-designed ideal do point, albeit in a general way, to a better product (Li et al. 2014).

Finding actual corporate case studies in the literature is difficult to do, but one industry-facing organization, ASTM, the American Society for Testing and Materials (Committee E-18 on Sensory Analysis), has created recommended practices for the JAR scale. That such attention is paid to the topic of standardizing the JAR scale attests to its practical use and importance for product design. The JAR scale is used to guide product redesign, with the directional referring to changes in the sensory attributes of the product being tested. In contrast, there seems to be no clear literature about the practical, industrial use of the self-designed ideal to guide product design, even though the approach has been around for more than 46 years (Drewnowski and Moskowitz 1985; Moskowitz 1972).

Let us assume that we can, in some way, predict the liking of the product which is perfect on the JAR scale, or which “delivers” the sensory profile of the self-designed ideal. Are there data showing that this product expected to perform in an optimal way, if we were able to estimate the liking of this product? Keep in mind that the JAR scale and the self-designed ideal work only with sensory attributes, and do not involve the key evaluative criteria of overall liking, or likelihood to purchase. In a study of pizza, with all ingredients disguised to maintain corporate confidentiality, Moskowitz demonstrated through modeling that one could create a set of equations

relating formula variables to liking, to sensory attribute levels, and to JAR scales. Using the model, it was possible to set the JAR scale values all to 0 (no change required) or to set the sensory attributes to the level defined by the self-designed ideal. The results suggested that the products emerging from this exercise were not optimally acceptable, i.e., the formulations expected to generate JAR scale values of 0 (just about right), or to generate the self-designed ideal (Moskowitz 2001).

It may be that the JAR scale and self-designed ideal work for attributes which are not inherently hedonic, such as appearance, texture, and flavor attributes. For those nonjudgmental attributes, respondents have a sense of what they want. For many other attributes, such as salty, fatty, and so forth, the JAR scale and the ideal levels must be interpreted with a note of caution because the respondent often either never gets enough (natural flavor) or always has too much (e.g., fatty for a health-oriented food). Table 1 (section A) shows the results of a commercially funded study with

Table 1 Data from a 1993 commercially funded study on consumer responses to 11 prototypes of frankfurters, prior to the selection of one product to “go to market.” No prototype ever scores sufficiently “meaty,” whether in terms of the JAR scale or in terms of the self-designed ideal

A - Average JAR (just about right) scale results for 11 frankfurter prototypes					
	JAR	JAR	JAR	JAR	JAR
Code	Dark	Smoky	Meaty	Salty	Greasy
101	6	-1	-8	2	-1
102	7	-4	-10	1	-3
103	-17	-8	-16	2	-1
104	-10	-11	-16	2	1
105	-21	-16	-14	5	-2
106	1	-8	-11	4	5
107	31	-16	-13	-5	2
108	-4	-7	-10	-2	-4
109	-20	-13	-20	-7	-5
110	-29	-14	-17	-4	-1
111	-4	-8	-21	9	1
B - Minimum, Self-Designed Ideal, and Maximum for 11 frankfurter prototypes on nine 'sensory' attributes					
	Min	Ideal	Max		
Dark	19	45	67		
Long	26	67	74		
Processed App	57	30	69		
Aroma	41	46	54		
Smoke Flavor	31	36	51		
Meaty	34	76	55		
Salty	30	25	51		
Aftertaste	40	32	61		
Firm	29	46	60		
Juicy	42	57	60		
Greasy	21	13	34		

“frankfurters,” wherein the respondent rated the sensory intensity of attributes, the sensory ideal of the attribute, and the JAR scale for the attribute. We show the results for the 11 products. The important thing to note is that none of the 11 test samples ever scored high enough on the attribute of “meaty.” Of course, it might well be that these frankfurter prototypes were simply not sufficiently “meaty,” but we find similar types of failure to deliver on other attributes which are “hedonics in disguise,” such as “real chocolate flavor, etc.”

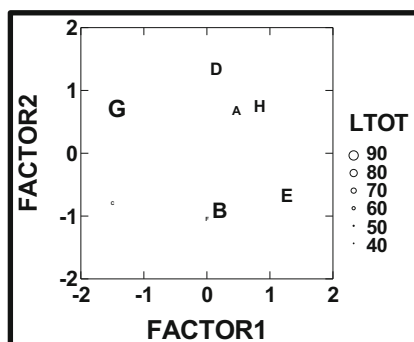
When we move to the self-designed ideal (Table 1, section B), we see the same type of problem emerging. One attribute, meaty, shows a self-designed idea outside the range of the levels achieved by the 11 prototypes. Again, the problem emerges, namely, “what should the developer do with these results?”

Multiple Test Stimuli: Mapping

Placing points on a geometrical space appeals to researchers. Whether the points define some type of function or the points define the location of an item in space, there is the perennial desire of a research to display data visually. Quite often, such displays reveal patterns that would be otherwise undetected. Mapping began with statisticians, who suggested that factor analysis, which reduces the dimensionality of a set of variables to a simpler set of orthogonal primaries, could be even more valuable when one plotted the stimuli as points in this orthogonal space, as Fig. 2 shows. The size of the letter is proportional to overall liking (LTOT).

When used for product design, maps reveal open areas, opportunities for new products. In today’s business parlance (2019), the term is “white space.” Author Moskowitz developed methods by which to identify the sensory profile of the products to be fit into this whole, using a method called “sensory-based engineering.” The approach used the coordinates of the map as independent variables and each of the sensory attributes and the rating of liking as separate dependent variables. The optimization routine identifies the coordinates in the factor space corresponding to the “best product” and then estimates the likely sensory profile of that best product (Moskowitz 1994).

Fig. 2 Example of a map, showing the location of products as letters, with the size of the letter proportion to product acceptance (LTOT = overall liking)



Multiple Test Stimuli: Response-Response Analysis

Regression analysis occupies the honorable position of being perhaps *the* statistical method to uncover so-called drivers of liking. The original use of regression analysis by statisticians involved the analysis of large data sets in order to identify which of the measured factors covary with a key evaluative criterion.

Focusing on food, the application translates into the simple problem of which of a variety of ingredients drives liking or which of a set of sensory attributes drive liking, respectively. The researcher assembles a set of products, either variants of each other in terms of ingredients/processes or of the same general type. The researcher instructs the panelists (consumers, experts) to rate the products on a set of scales (e.g., sensory perception of color, aroma, taste, mouthfeel) and instructs the consumers to rate the products on an evaluative criterion, e.g., “liking” or “purchase intent.” The latter evaluative ratings are obtained either from the same panelists providing the sensory ratings or from other panelists representing the ultimate consumer.

When the researcher works with a single set of products, whether or not these products are systematically related to each other by an underlying design, it is straightforward to plot the relation between acceptance (e.g., overall liking) on the ordinate and sensory attribute level on the abscissa. Figure 3 shows the results for the study of the frankfurters, from personal data collected by author Moskowitz in 1993.

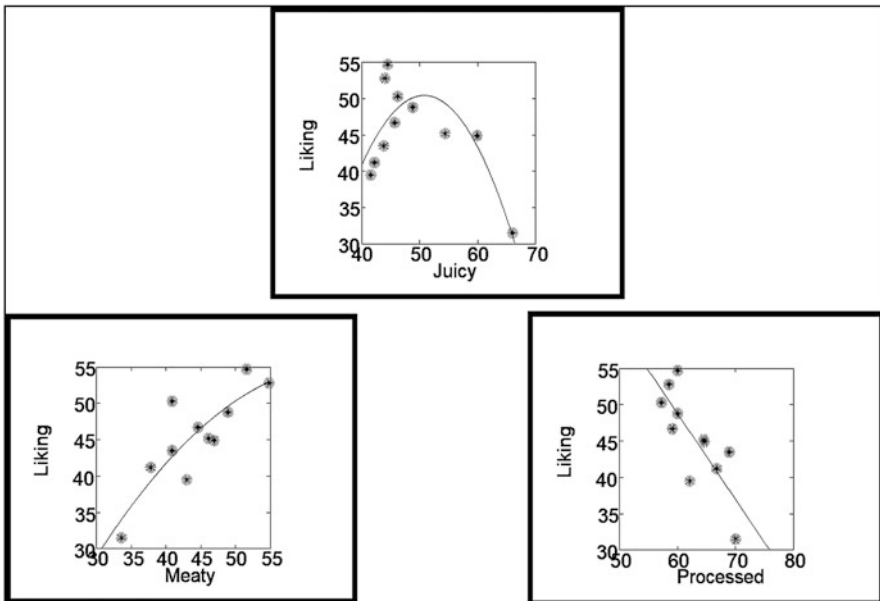


Fig. 3 The scatterplot relation between sensory attribute level and overall liking for the 11 frankfurters introduced in the previous sections. The filled star/circles correspond to the 11 samples. The statistical program (Systat) fits a quadratic function to the scatterplot

The abovementioned approach is called R-R analysis, or response-response analysis. We look for relations between two variables, neither of which is systematically varied. Rather, both variables emerge from the evaluation of the same 11 meat samples. We cannot discover “causality” but simply get a sense of which attribute covaries with liking.

Systematics: Creating and Using Psychophysical Curves

The new psychophysics has had a direct impact on the world of product design, perhaps one that was anticipated. Over the earlier decades of the twentieth century, researchers began to explore the use of people as “measuring instruments.” For food, this systematized effort began in earnest with ratings of liking of different products, with such ratings being analyzed for differences (Peryam and Pilgrim 1957). Some researchers realized, however, that they had a tool by which to understand how ingredients drove responses, whether the response is the perceived sensory intensity of the food or beverage (e.g., the sweetness of cola) or the degree of liking. Psychophysicists, specializing in the study of the relations between sensory magnitude and physical intensity, soon began to contribute to this effort, especially with simple systems, such as colas and some foods (Moskowitz et al. 1979). What is important to keep in mind is that these curves provided foundational knowledge. It was quickly discovered that the same percent change in ingredients could very well produce radically different perceived changes. Doubling the concentration of a flavor ingredient, for example, was seen to be less effective than doubling the concentration of sugar for the same food. We will elaborate this type of thinking below, when we deal with stimulus-response analysis and response surface designs.

Stimulus-Response Analysis

Product design becomes far more powerful when we move from testing one product or several unrelated products to a set of prototypes that are systematically varied. The original thinking comes from both statisticians who promoted the idea of DOE (design of experiments) and from psychophysicists who promoted the idea of uncovering lawful relations between physical stimuli and subjective responses.

Beyond the design of experiments, the creation of the systematically varied prototypes lies on an entire body of statistics known as regression analysis or curve fitting. With regression analysis, one discovers how a physical ingredient or set of ingredients drives a response, the most important response being overall liking. Furthermore, with nonlinear regression analysis, it becomes easier to uncover optimal or most highly liked product formulations within the range of prototypes tested, but a formulation not necessarily one of the prototypes created.

Figure 4 shows a schematic example of the approach. The regression modeling typically creates either a linear plot, described by the equation and shown as

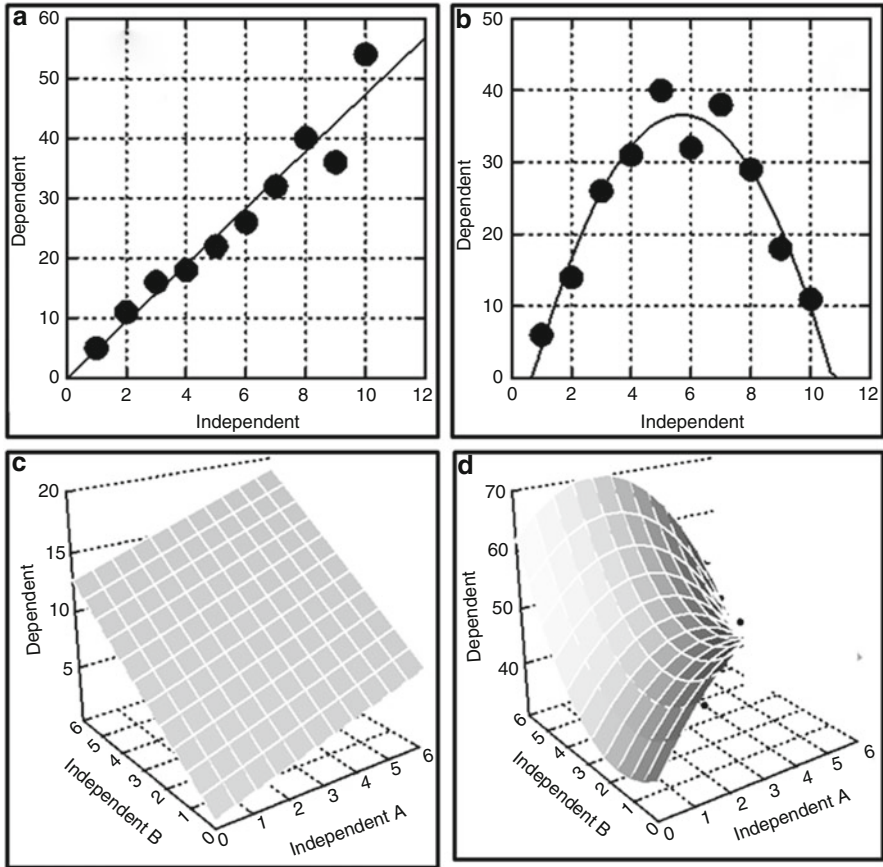


Fig. 4 Schematic example of scatterplots showing the relation between a dependent variable (ordinate) and either one independent variable (panels A and B) or two independent variables (panels C and D). The plots represent the type of relations one might observe when the dependent variable (ordinate) continues to increase linearly with increases in the independent variable (panel A, panel C) or when the dependent maximizes at some intermediate point of one or both independent variables (panel B, panel D)

examples (left panels) or described by a quadratic equation, in Fig. 4 (right panels). The plot can either reflect a one-ingredient system (Panel A, Panel B) or a two-or-more ingredient system (Panel C, Panel D). We present a visual only of the two-ingredient system, but the mathematical modeling can accommodate many more independent variables, even when we cannot easily visualize the model.

Regression modeling plays an important role in product design for at least two reasons.

1. Regression provides insights into what might be important, giving specific, testable direction for product design. The developer creates a quantitative

structure to discover what operationally varied factor might be important, rapidly providing insights that could not be obtained were the effort to be focused on one product.

2. The second, perhaps more important, is that regression analysis forces the necessary shift from focusing on one product to focusing on many products. One soon realizes the futility of efforts to understand the drivers of liking by relating the rating of liking for one product to the rating of sensory attributes for the same product, using as inputs, both the sensory and the liking ratings assigned to a single product by many respondents. The reliance on one product alone produces a fallacious approach, confusing the “noise or variability-based” information in the variation of responses to one product with the “signal-based” information from the variation of responses to many products. As the intellectual development of the field proceeded, this change in focus, from the study of variability to the study of patterns, would inevitably lead to more powerful tools.

Expectations

Another methodological breakthrough in the development of consumer methodology for food and drinks was the application of expectation theory and expectation methods. Perhaps the lead researcher in this area has been Armand Cardello from the US Army Natick Laboratories. Cardello (2007) lays out the history of expectation theory and research. Like some other methods in this chapter, expectation theory and methods developed within psychology in the earlier part of the twentieth century. Broadly speaking, expectations refer to the anticipation that something will occur; when applied to foods and drinks, expectations refer to the anticipation that a product will contain an attribute or result in a consequence. Early discussion of expectations focused in its role in human behavior, especially motivation and cognition. Expectation was applied to consumer research in the 1970s in the context of models of consumer satisfaction and service quality. This satisfaction work led Cardello and colleagues to apply the expectation model to product satisfaction (Cardello et al. 1985; Cardello and Sawyer 1992).

Expectation research uses the model of confirmation/disconfirmation in which a person’s expectations are either met (confirmation) or not met (disconfirmation), resulting in a product producing satisfaction or rejection. Research has shown that the acceptance rating of a product often moves in the direction of its expectation, referred to as assimilation. If you expect a product to be good, you will rate it higher than if you expected a lesser rating. This is important for product design, product advertising, and product success, because final acceptance by the consumer is due to the actual product attributes but also to the expectation of those attributes. Product liking is due to more than the physical make-up of the product, as noted above. In addition to the assimilation model of expectations, there are also other models including (1) generalized negativity, (2) contrast, and (3) assimilation-contrast. The contrast model is one that is observed when a product

falls much below its expectation, leading to rejection. Setting product expectations too high can lead to product rejection.

Conclusions

As we reach the end of this history, and this writing (2019), we are approaching the greater part of a century since the first efforts at applying knowledge about the human senses have been applied in a significant way to the development of foods. What then have we learned in the past century? What has really transpired, and how have we developed? There are some key trends that we should note.

1. *Descriptive analysis*: Research revealed that for easily understood attributes, the ability of consumers to detect and report the magnitude of changes was equivalent to that of experts, at least for sensory attributes that were obvious, and did not need explanation to be recognized before being evaluated in an experiment (Moskowitz 1997; Ares and Varela 2017). Research has shown that it's not all about experts, with consumers relegated to relatively blunt instruments whose only ability was to respond with "like, neutral, or like." Rather, one should look at experts as trained to recognize notes and to describe whether this description is relevant to product design or simply limited to quality control. It is important, to note, however, that consumers, even untrained consumers, can do a reasonably "good job" evaluating sensory attributes that are easy to understand and with which they have had everyday experience. However, some commercial research heavily depends on trained panels for quality control and other functions.

In a broader sense, trained panel work has not delivered what people thought it would deliver. There was a sense, perhaps not well stated, that if we understand the sensory properties of the food, we could formulate newer and better products. Unfortunately, the promise of descriptive was never fulfilled in an operationally defined way. That is, there seems to be no simple relation between how a product is described and what the product developer must do to change the product.

2. *Product testing*: Testing products has survived and flourished. Today, more than ever, corporations depend upon the scores in so-called product tests to move forward with a prototype toward market, to modify the prototype, or sometimes to just "kill" the project. Testing, in fact, of all the contributions from human research, has been the most absolutely robust, perhaps because it is structured and well-choreographed, admits to "best practices," and can be supported by numbers, by statistics. All three of these reasons make it easy to adopt testing as a standard procedure in corporate work.

But a number of things have changed in product testing. The product evaluations are typically done with "target consumers," i.e., with consumers who are representative consumers of the products. They are less frequently done with

convenience panels, i.e., panelists who are company employees and available for testing. For larger companies, product testing is often done in multiple countries, without the assumption that one test in one country guarantees global success. What remains to be done is to better design these cross-cultural tests, which remain a challenge, as discussed at a Pangborn Sensory Science Symposium (Goldman 2006). Another change in product testing is the growth of the field of sensometrics, with much more advanced statistical designs for data collection and analysis. A final change is the growing appreciation of the context in which product tests are conducted, with a growing use of home testing and non-laboratory testing. Of great interest here is the growth of virtual reality as an alternative to changing actual testing locations (Meiselman 2019).

3. *Experimental design*: Experimental design of products has had its ups and downs, due in great measure both to the benefits it provides and to the effort and cost it demands. Experimental design forces the developer to create products, an effort often resisted because it requires time, investment, and effort, all three in a world which seeks success using faster and less expensive methods. Those are the downsides, which stop experimental design in its tracks and limit its true value. We can expect more experimental design work, however, when the methods become less expensive, faster, and obvious, rather than seeming esoteric and unapproachable.
4. *Expanding the field – sensory becomes sensory and consumer*: In the early 1970s, the Institute of Food Technologists formed the Sensory Evaluation Division (SED). At that time, the world of human food research as we think of it comprised the so-called sensory researchers in laboratories. There were some researchers such as author Meiselman who campaigned for the broader study of food habits (Meiselman 1992), but the majority of research facilities and research focus remained steadfastly on what we today would call sensory issues, with foods as the primary focus and the person as a secondary, convenient instrument on par with instruments but of course an instrument which “evaluated” as well (good versus bad, etc.). Over the decades, however, the world of human food research expanded its borders, incorporating market researchers and anthropologists. The IFT later changed its name to the Sensory and Consumer Sciences Division to recognize the importance of consumers.
5. *Data analysis*: The advent of the computer has brought with it many methods. It is hard to know which methods have lasting impact. There are those which appear at one or another conference, and have “staying power” such as temporal dominance as a research tool or conjoint measurement for messages as research and as a development tool. There are also methods that are accepted, but their utility is less clear, such as mean drop analysis, a method of cross tabulation of one product to find out what are the important attributes. We can be reasonably assured that the continually increasing group of young, sophisticated, and motivated researchers will continue to introduce new data analytic methods at today’s pace and no doubt increase the pace in the future.

References

- Amerine, M. A., Pangborn, R. M., & Roessler, E. B. (1965). *Principles of sensory evaluation of food*. New York/London: Academic.
- Ares, G., & Varela, P. (2017). Trained vs. consumer panels for analytical testing: Fueling a long lasting debate in the field. *Food Quality and Preference*, *61*, 79–86.
- Bearden, W. O., Netemeyer, R. G., & Mobley, M. F. (1993). *Handbook of marketing scales* (1st ed.). London: Sage.
- Bell, B. L., Oshinsky, N. S., & Wolfson, J. (1965). *Food acceptance and preference research: An annotated bibliography* (Technical Report EPT-5, 58 pp). Natick: US Army Natick Laboratories.
- Benjamin, L. T., Jr. (1988). *A history of psychology: Original sources and contemporary research*. New York: McGraw-Hill.
- Box, G. E., Hunter, J. S., & Hunter, W. G. (2005). *Statistics for experimenters: Design, innovation, and discovery* (Vol. 2). New York: Wiley-Interscience.
- Cairncross, S. E., & Sjöström, L. B. (1950). Flavor profiles: A new approach to flavor problems. *Food Technology*, *4*, 308–311.
- Cardello, A. V. (2007). Measuring consumer expectations to improve food product development. In H. MacFie (Ed.), *Consumer-led product development* (pp. 223–261). Cambridge: Woodhead.
- Cardello, A. V., & Sawyer, F. M. (1992). The effects of disconfirmed consumer expectations on food acceptability. *Journal of Sensory Studies*, *7*, 253–277.
- Cardello, A. V., Maller, O., Bloom-Masor, H., Dubose, C., & Edelman, B. (1985). Role of consumer expectancies in the acceptance of novel foods. *Journal of Food Science*, *50*, 1701–1714, 1718.
- Caul, J. F. (1957). The profile method of flavor analysis. In *Advances in food research* (Vol. 7, pp. 1–40). New York: Academic.
- Civille, G. V., & Szczesniak, A. S. (1973). Guidelines to training a texture profile panel. *Journal of Texture Studies*, *4*(2), 204–223. <https://doi.org/10.1111/j.1745-4603>.
- Cochran, W. G. (1950). *Experimental design*. New York: Wiley.
- Costa, A. I., & Jongen, W. M. F. (2006). New insights into consumer-led food product development. *Trends in Food Science & Technology*, *17*(8), 457–465.
- de Graaf, C., Kramer, F. M., Meiselman, H. L., Leshner, L. L., Baker-Fulco, C., Hirsch, E. S., & Warber, J. (2005). Research review: Food acceptability in field studies with US army men and women: Relationship with food intake and food choice after repeated exposures. *Appetite*, *44*(2005), 23–31.
- Dove, W. F. (1947). Food acceptability: Its determination and evaluation. *Food Technology*, *1*, 39–50.
- Drewnowski, A. (1993). Studies of food quality and preference: The contribution of Rose Marie Pangborn. *Food Quality and Preference*, *4*(1-2), 3–9.
- Drewnowski, A., & Moskowitz, H. R. (1985). Sensory characteristics of foods: New evaluation techniques. *The American Journal of Clinical Nutrition*, *42*(5), 924–931.
- Ferber, R. (1974). *Handbook of marketing research*. New York: McGraw Hill.
- Goldman, A. (2006). Workshop summary: Cross-cultural sensory and consumer studies. *Food Quality and Preference*, *17*(7), 646–649.
- Jones, L. V., Peryam, D. R., & Thurstone, L. L. (1955). Development of a scale for measuring soldiers' food preferences. *Food Research*, *20*(5), 512–520.
- Land, D. G. (1988). Chapter 1. Roland Harper – Pioneer food psychologist. In D. M. H. Thomson (Ed.), *Food acceptability* (pp. 3–10). London: Elsevier Applied Science.
- Li, B., Hayes, J. E., & Ziegler, G. R. (2014). Just-about-right and ideal scaling provide similar insights into the influence of sensory attributes on liking. *Food Quality and Preference*, *37*, 71–78.
- Meilgaard, M., Civille, G. V., & Carr, B. T. (1999). *The Spectrum descriptive analysis method sensory evaluation techniques* (3rd ed.). Boca Raton: CRC Press.
- Meiselman, H. L. (1992). *Methodology and theory in human eating research*. *Appetite*, *19*, 49–55.

- Meiselman, H. L. (2019). *Context: The effects of environment on product design and evaluation*. Kindlington: Woodhead Publishing.
- Meiselman, H. L., & Schutz, H. G. (2003). History of food acceptance research in the US Army. *Appetite, 40*, 199–216.
- Moskowitz, H. R. (1972). Subjective ideals and sensory optimization in evaluating perceptual dimensions in food. *Journal of Applied Psychology, 56*(1), 60–66.
- Moskowitz, H. R. (1994). *Food concepts and products: Just-in-time development*. Trumbull: Food and Nutrition Press.
- Moskowitz, H. R. (1997). A reply note on experts versus consumers: A comparison. *Journal of Sensory Studies, 12*(2), 155–168.
- Moskowitz, H. R. (2001). Sensory directionals for pizza: A deeper analysis. *Journal of Sensory Studies, 16*(6), 583–600.
- Moskowitz, H. R., & Wehrly, T. (1972). Economic applications of sweetness scales. *Journal of Food Science, 37*(3), 411–415.
- Moskowitz, H. R., Wolfe, K., & Beck, C. (1979). Sweetness and acceptance optimization in cola flavored beverages using combinations of artificial sweeteners – A psychophysical approach. *Journal of Food Quality, 2*(1), 17–26.
- Narayanan, P., Chinnasamy, B., Jin, L., & Clark, S. (2014). Use of just-about-right scales and penalty analysis to determine appropriate concentrations of stevia sweeteners for vanilla yogurt. *Journal of Dairy Science, 97*(6), 3262–3272.
- Pangborn, R. M., Simone, M., & Nickerson, T. A. (1957). The influence of sugar in ice cream. I. Consumer preference for vanilla ice cream. *Food Technology, 11*, 679–682.
- Peryam, D. R., & Girardot, N. F. (1952). Advanced taste-test method. *Journal of Food Engineering, 24*, 58–64, see also page 194.
- Peryam, D. R., & Pilgrim, F. J. (1957). Hedonic scale method of measuring food preferences. *Food Technology, 11*(9), 9–14.
- Schutz, H. G., & Pilgrim, F. J. (1957a). Differential sensitivity in gustation. *Journal of Experimental Psychology, 44*(1), 41–48.
- Schutz, H. G., & Pilgrim, F. J. (1957b). Sweetness of various compounds and its measurement. *Food Research, 22*(2), 206–213.
- Stevens, S. S. (1961). To honor Fechner and repeal his law. *Science, 133*, 80–86.
- Stevens, S. S. (1975). *Psychophysics: An introduction to its perceptual, neural and social prospects*. New York: Wiley.
- Stone, H., Sidel, J. L., Oliver, S., Woolsey, A., & Singleton, R. C. (1974). Sensory analysis by quantitative descriptive analysis. *Food Technology, 8*, 24–32.
- Valdes, R. M., & Roessler, E. B. (1956). Consumer survey on the dessert quality of canned apricots. *Food Technology, 10*, 481–486.
- Vanaja, K., & Shobha Rani, R. H. (2007). Design of experiments: Concept and applications of Plackett Burman design. *Clinical Research and Regulatory Affairs, 24*(1), 1–23.
- Wells, W. D. (1975). Psychographics: A critical review. *Journal of Marketing Research, 12*, 196–213.