THE EFFECT OF CERTAIN PSYCHOLOGICAL AND ENVIRONMENTAL FACTORS ON FLAVOR PREFERENCE TESTING

by

LYNNE ARCHER PETTIT

A THESIS

submitted to

OREGON STATE COLLEGE

in partial fulfillment of
the requirements for the
degree of

DOCTOR OF PHILOSOPHY

June 1956

APPROVED:

Professor of Food Technology

In Charge of Major

Head of Department of Food Technology

Chairman of School Graduate Committee

Dean of Graduate School

Date thesis is presented May 1, 1956

Typed by Alura Paul

ACKNOWLEDGMENT

Particular appreciation is extended to Dr. Jerome C. R. Li for his inspiration and counsel in working out the designs for these experiments and for his guidance in the statistical analyses; to Dr. J. W. Sherburne for his discussions and suggestions concerning psychological interpretations; and to Mrs. Lois Sather and her assistants for their tireless efforts in the serving of samples and the collecting of data.

TABLE OF CONTENTS

	Page
Introduction	1
Definition of flavor	1
Measurement of flavor	4
Flavor preference testing	6
Factors involved in flavor preference	8
Review of Literature	10
Methodology	10
Training, experience, age	15
Suggestion, knowledge, and related factors	20
Physiology, sensitivity, sex	24
Rinsing, swallowing, fatigue, hunger	28
Color, appearance, sound	31
Experimental Approach	34
Objectives	34
Method of evaluation	35
Scoring system	35
Flavor panels	36
Product	36
Effect of the Number of Samples Presented Simultaneously	38
Procedure	38
Analyses and results	41
Summary	43
Effect of Location of the Test and of Sound Accompanying the Test	45
Procedure	45
Locations	46
Sound	47
Analyses and results	49
Summary	52
Effect of Information Possessed by the Taster	53
Procedure	53
Analyses and results	57
Summary	62

Table of Contents - Continued

Page

Effect of Rinsing Between Tastes..... 64

Procedure..... 64
Analysis and results..... 69
Summary..... 70

Effect of Swallowing the Product..... 72

Procedure..... 72
Analyses and results..... 76
Summary..... 77

Effect of the Quantity of Sample Consumed.... 79

Procedure..... 79
Analyses and results..... 83
Summary..... 85

Summary and Conclusions..... 87

Bibliography..... 89

LIST OF TABLES

Table		Page
1	Effect of number of samples presented simultaneously.....	42
2	Effect of number of samples -- mean scores of duplicate samples to the reference	43
3	Effect of location and sound -- duplicate Sample A.....	50
4	Effect of location and sound -- duplicate Sample C.....	50
5	Effect of Location and Sound -- Sample B with Reference A.....	51
6	Effect of Location and Sound -- Sample B with Reference C.....	52
7	Effect of information re: raw product quality.....	59
8	Effect of information re: chemical treatment	60
9	Effect of information re: exposure to rays	60
10	Effect of information re: addition of MSG	60
11	Effect of information re: addition of salt	61
12	Effect of information re: addition of sugar	61
13	Effect of information re: addition of lemon juice.....	61
14	Effect of information -- mean scores of duplicate samples to the reference	62
15	Effect of rinsing between tastes.....	70
16	Effect of rinsing -- mean scores of coded samples.....	70
17	Effect of swallowing the product.....	76

List of Tables - Continued

Table		Page
18	Effect of swallowing -- mean scores of coded samples.....	77
19	Effect of quantity of sample consumed -- Test A--samples identical to the reference.....	84
20	Effect of quantity of sample consumed -- Test B--samples different from the reference.....	84
21	Effect of quantity of sample consumed -- mean scores of coded samples.....	84

THE EFFECT OF CERTAIN PSYCHOLOGICAL AND ENVIRONMENTAL FACTORS ON FLAVOR PREFERENCE TESTING

INTRODUCTION

This investigation into factors involved in flavor was prompted by observations that the meaning of flavor was not clearly prescribed among food technologists, and that the problems connected with the determination of flavor were essentially those problems involved in the measurement of human judgment. The diversity of methods utilized for the evaluation of flavor suggested that flavor is complex, to be approached only through various pathways emanating from incomplotely understood human responses.

Definition of Flavor

The meaning of the word "flavor" would seem to be self-evident in view of its long standing association with foodstuffs. Nevertheless, all workers in the field of assessment of food quality are not in clear agreement. Crocker (12) has called it the "voice of food," affecting "every sense to some extent, but principally the so-called 'chemical' senses of taste and smell.....To understand flavor completely," he says, "we must consider not only its chemistry and physics but also the physiology and psychology of its perception and of our reactions to it."

He defines flavors (10) as "substances used for stimulation of the so-called 'chemical senses,' those of taste, mouth-feel, and odor."

Kirchner (34) says that the popular conception of flavor is that it is synonymous with taste. Possibly this is because there is no specific verb to describe the act of determining the flavor of a substance. "When we taste," he says, "we are actually judging the flavor, for flavor is an elusive blend of sensations. It includes not only taste but also the sense of smell and touch."

Sjostrom (55) describes flavor as "that property of a food or beverage which excites mainly the senses of taste and smell. There are four fundamental taste sensations which are experienced when a food is placed on the tongue: sweetness, sourness, saltiness, and bitterness."

A newspaper reporter (28) recently wrote: "Flavors, of course, are chemical compounds. But the problem of studying them.....also embraces psychology, physiology and philosophy."

Kramer (35) suggests that "since flavor is a physiological sensation, it is by definition subjective, and therefore, not subject to objective measurement."

Webster's new international dictionary (58) defines flavor as that quality of anything which affects the smell, the taste or gratifies the palate.

These various explanations imply that flavor may be an intangible without standardized meaning. This makes it necessary for an investigator in this field to establish for himself a definition of what is meant by flavor before proceeding to probe into its manifestations through scientific research. The lack of unity among the foregoing explanations leaves adequate opening for added opinions provided they are based on thoughtful and reasonable considerations.

It appears to this writer that to hold that flavor is either a physical property of chemical compounds would be to ignore that no substitute for identification of flavor has as yet been found for the human, or animal organism, or to ignore that flavor itself appears as a result rather than a cause. Flavor may be more precisely defined as a mental evaluation of food-invoked sensory stimulations. Food, itself, does not actually possess flavor. It provides physical and chemical changes to which the organism can be sensitive. Interpretive judgment expressed as flavor arises within the organism in consequence of such changes. Food consists of chemical compounds which may act to initiate flavor response, but which themselves are not flavor.

Flavor, as I view it, is a perceptive judgment arising from the integration of sensory stimuli with other factors

of mental processes to produce meaning. These other factors may include such psychodynamic characteristics as memory, experience, attitudes and expectations, as well as the influence of environment or society. In this light, flavor is more properly considered an attribute of the taster rather than an attribute of food. That such a consideration has in fact been involved in the development of flavor methodology is supported by examination of the methods used to evaluate flavor. These methods are simply devices to obtain and interpret human judgment.

Food, of course, has a very essential role in the creation of flavor. It provides the sensory stimuli which may then be channeled through the brain for integration. But food remains an initiating stimulus, not an effect, while flavor appears as an effect. Food can be examined directly as a component of flavor or as an assembly of chemical or physical identities capable of bringing about stimulation of the senses. Flavor itself can be examined only in the mental processes of the consumer of food.

Measurement of Flavor

Methods of measuring flavor are systematic efforts to obtain and evaluate human judgments. Methods differ in specific objectives, in procedures, in systems of interpreting results, in efficiency and economy. The

objective for the test is the most critical consideration in the selection of a test method. Responses of tasters--or judgments--which are the end product of a flavor test may be obtained to serve several specialized purposes in flavor evaluation. These purposes furnish a convenient basis for classifying current flavor testing procedures into the following groups:

1. Difference tests--which are intended to ascertain only the existence of a difference in flavor between foods. Examples of these are the triangular test, the paired difference test, and the reference difference test.

2. Preference tests--which are value comparisons in terms of taster preference between food samples. Ranking test, paired preference test, and reference-type preference test are used for this purpose.

3. Acceptance tests--which are aimed to determine the position of desirability of a food product. Among the devices to examine this elusive aspect of flavor are rating scales (numerical and word), hedonic scales, and consumer polls.

4. Description or identity tests--which strive to identify flavor in terms of precise words, descriptive phrases, or attributes. Among these are the flavor profile, identification tests, and description scales.

Flavor Preference Testing

Opinions as to acceptability of food products are not always made in an absolute sense with respect to liking or disliking or rating. They may be made as comparisons only of one food against another. These are preference judgments, expressed without specific regard to position on a scale of acceptance values but only as to whether a product occupies a preferred or not preferred position in a comparison situation.

Foods are most often compared in pairs for this kind of flavor evaluation. One of a pair may be designated as the reference, or no reference may be specified in the pair. It is possible that preference judgment may be expressed in the presence of only one food sample, relying upon memory or on a mental standard to furnish the other sample for the comparison. This, of course, is not without its hazards of unreliable or uncertain reference points. Several foods may be ranked in the order of preference to obtain an array of preferences at one time.

Preference testing of food products has a value of another nature than that of difference testing. It is thought by some to be more productive than pure difference work because it adds a qualitative interpretation to the demonstration of differences between food flavors. That is to say, it not only has the ability to

show the existence of meaningful differences--for when a preference is shown it signifies the existence of a difference--but it also gives a meaning in terms of preferred or not preferred to the differences thus revealed. Flavor difference testing cannot yield this meaning without supplemental help. Differences, of course, can exist without the presence of preferences, but such differences in flavor pose a serious question as to what they mean. Interpretation of the meaning of flavor differences is a most practical consideration to the producer and manufacturer of food products. These people are generally, therefore, more attracted to flavor evaluation methods which reflect preferences or acceptance in food products.

Inasmuch as preferences for flavor are expressed opinions from the minds of persons, it may be expected that a question arises as to how many opinions constitute a valid basis for establishment of trends in preferences. An immediate forthright answer to this cannot be given, as it brings up the entire area of population sampling with its many dependent variables. The practice in preference testing is to obtain judgments from as many individuals as possible, then to recognize that conclusions drawn from them may reflect only the particular group, or the particular population of which the judging group might have been an adequate and representative sample. The existence

of a population "expert" in preference has yet to be demonstrated--it may never be. Every opinion obtained on preference, barring subterfuge, must be accepted at face value as a straightforward and genuine judgment--biased or not--entitled to its share of influence in any assembly of results. This may further emphasize the need for large numbers of judges in order to guard against undue influence by individuals or by small groups in predicting the valid preference ratings of consumers at large.

In spite of what may appear to be subjective handicaps and hazards as an analytical tool, flavor preference testing is nevertheless a valued aid in determining the effects of changes in food flavors. If proper consideration is given to factors which may influence its results and to adequate statistical procedures for its experimental use and interpretation, flavor preference testing can provide practical information concerning the potentials of food.

Factors Involved in Flavor Preference

Inasmuch as preference consists of opinion and flavor preference testing consists of the systematic collection of opinions, it can be seen that, as in other kinds of opinions, the possibilities for influence are many. The food itself, of course, exerts a real influence, but other less tangible forces sometimes appear to affect the

judgments. On occasion, it appears that factors are given credit by observation and logic for having an influence on consumer opinion which influence may not have a basis in demonstrated fact.

To enumerate a list of the many factors which could be involved in flavor preference judgment would be to recite all of the aspects of human perception. These would include each of the human sensory receptors, the nervous system and bodily chemistry, as well as musculature, reacting mechanisms, acuity of receptor cells, conductivity of the nerve cells, and other biological factors. Psychological and sociological factors of various kinds should be considered for their potential influence on the final meaning produced from the transmitted sensory stimuli. Influencing factors may include experience, expectations, fatigue, mental attitudes, memory, methods of testing, surroundings, suggestions, needs, and others, the omission of which does not imply a lack of possible importance. Such a partial recital of those influences which might be brought into play in forming a preference judgment serves to demonstrate the dangers of over-simplifying the meaning of flavor. It emphasizes also that flavor preference testing should be viewed as no less than an exploration into a complex pattern of human behavior.

REVIEW OF LITERATURE

Methodology

Literature in the field of flavor in general is widely scattered, diversified in nature, and largely uncorrelated. Literature dealing specifically with flavor preference is relatively scarce. The great dispersion of objectives among test methods is noted by investigators in pointing out the problems of obtaining agreement on the approach to flavor evaluation. Efforts at standardization have been dealt with cautiously, even with reluctance by some. The general status of the field of flavor testing is frequently referred to as new, in a state of flux, or growing. Proposed new methods of sensory evaluation appear often, while older methods are adapted to new variations and to new procedures of interpretation. Peryam (49), in 1953, stated that "Certainly the field is not ready for rigid standardization; perhaps the absence of attempts at standardization represents no more than proper scientific caution." At the same time, he was one of the earliest to spearhead a drive toward standardization of some of the existing procedures in promoting specifications for the sensory evaluation of military subsistence items.

Peryam classified flavor testing methods into five types: (a) The "palatability rating" test, palatability

being defined as "the resultant of sensory properties which tend to make a food pleasing and acceptable to the consumer." Scales employing the terms good, fair, poor, etc are utilized for this. (b) The flavor identity test, with its expressed purpose of determining the extent to which a single product sample is identified as possessing a specified flavor. Trained tasters may be used here. (c) The "paired-comparison" technique whose stated purpose is to determine whether a sample is equal or superior to a standard sample on the basis of "some defined criterion, such as palatability, flavor intensity, or flavor purity." (d) The "triangle test," familiar to many as a discrimination test calling for the determination of flavor difference between two sample products by selection of the odd sample (or pair of samples) in a group of three. (e) The "dilution number test," being described as a test to determine "that amount of a sample, expressed as a percentage of the mixture, which may be mixed with standard material without making the mixture noticeably different from the standard itself."

These types may be further generalized into two main groups, described as discrimination tests for measuring sensory differences and rating scales for measuring subjective impressions.

Harper (24) classifies subjective judgments into two

main types, namely comparative judgments and judgments made without a comparison standard, sometimes called "absolute judgments." He points out that this latter is probably a misnomer as it is likely that there is always some sort of mental standard involved.

Ehrenberg and Shewan (17) have proposed an "objective" procedure for the development of laboratory sensory measurements based upon the use of descriptive terms. They suggest the following: "1. Take a given limited array of sensory phenomena (such as flavors and odors of fish). 2. Deliberately devise descriptions for these phenomena, and possibly attach numerical symbols to them for greater convenience. 3. Train people to use the same description for any given phenomenon. 4. Observe empirically whether the descriptions are used consistently, and discover the factors which.....will increase this consistency." It is clearly admitted that in order to approach the goal of objectivity, descriptions for such testing must be cautiously selected and must separate themselves from product desirability, personal preferences and consumer reactions. The difficulties of selecting appropriate descriptive terms generally are discussed by Crocker (9) and by Wagner (57). Ehrenberg and Shewan admit that their procedure rests upon the selection of objective descriptions.

Kramer (35) divides sensory testing into two divisions,

namely difference tests and preference tests. He states that the former are used principally as laboratory tools while the latter are involved in consumer-survey panels. He classifies taste testing designs into three types, paired comparisons, triangular comparisons, and multiple comparisons. The paired design is considered the simplest, the triangular more efficient due to reduced opportunity for correct guessing, and the multiple sample method more effective for less homogeneous and larger-unit products. Paired comparison is said to find its main application in consumer survey studies.

A single-sample method for detecting flavor differences is described by Neubert and Carter (48) in connection with work on apple juice. In this method samples were presented one at a time, at not less than three-hour intervals, and tasters were asked to indicate whether or not they detected a foreign flavor. Results of this method in detecting foreign flavor in insecticide treated apples and juice derived therefrom were more conclusive than results from the triangular difference method.

According to Dawson and Harris (14), methods of measuring sensory differences in food quality are of six basic types: (a) Scoring, which involves the use of numbers to designate values of particular characteristics of a food; (b) Ranking, in which food samples are placed

in decreasing or increasing order with respect to some characteristic; (c) Paired comparison which permits a judge to select one sample of two for preponderance of a given characteristic; (d) Triangle or triple comparison which allows a judge to choose a different sample from among three; (e) Dilution test, in which judges determine the smallest amount of an unknown that can be detected when mixed with a standard food material; (f) Descriptive terms which are used to define a food in terms of characteristics. This method is believed by Dawson and Harris to be difficult to analyze and report.

Boggs and Hanson (7) draw a careful distinction between sensory preference tests and sensory difference tests. They state that the object of a consumer acceptance or preference test is to determine what a representative population prefers whereas the difference test aims at measuring differences in samples irrespective of preference. Preferences for samples of food may vary in various locations, but sensory difference tests, they believe, should yield the same results within experimental limits no matter where they are conducted. In this light, they classify preference tests as "subjective" measurements and difference tests as "objective" measurements.

The differences between laboratory preference panels of fifty to seventy-five persons and household consumer

panels of approximately six hundred were studied by Miller, Nair and Harriman (44). Two samples of soup were mailed to consumers with instructions to serve one of them at meal time on each of two different days, then to record their judgments on a questionnaire. Laboratory panels compared the same products at one sitting in a paired preference test. There was general agreement in the preferences obtained by the two kinds of panels. The findings suggested that with this food product the panels tended to show a preference bias for the sample tasted first. Two consumer panels comparing the same samples, differing only in flavor, did not express the same preference. Two consumer panels comparing the same samples differing distinctly in appearance chose the same samples as preferred. A consumer panel which found a preference on the first test between two samples indicated no preference in the second test. These findings tended to strengthen confidence in laboratory panels but served to emphasize hazards in attempting to predict general consumer preference from panel studies.

Training, Experience, Age

Holm and Trolle (25) conducted an extended series of triangular taste tests, involving several thousand taste judgments with a panel of seventy people who evaluated

flavor differences in several types of beer. The objective was to obtain from this group a panel of "expert tasters," based upon those who consistently obtained the highest number of correct identifications of different samples. Results were evaluated with relation to occupation of the taster, previous experience in tasting, and practice.

Among these people, the highest percentages of correct discriminations of beer flavor differences were made by brewers and chemists, with engineers, architects, and women (as a group) making lesser percentages of correct determinations. The lowest percentages of correct answers were consistently obtained by foremen and supervisors.

Individuals with previous beer tasting experience obtained consistently higher percentages of correct identifications than the inexperienced tasters. Practice in tasting, as reflected by results in the latter portion of the tests as compared to the early tests, made no difference in the performance of the less discriminating tasters. On the other hand, the more discriminating tasters--the "experts" as they were designated--showed a definite trend toward improved discrimination in the latter part of the series. The number of correct identifications made by them as the tests reached the final stages increased, which was believed to reflect an ability of the discriminating taster to learn better by experience than the less

discriminating taster.

Those investigators concluded that small panels of discriminating tasters could be effectively used for determining the existence of differences in beer samples, but that quality evaluations should be obtained only by much larger groups in view of demonstrated variations in individual taste discrimination and in view of variable taste performance arising from previous experience and training.

Harper (24) emphasizes the effect of training and experience in refining human judgments, but draws a careful distinction between the two. "Training," he says, "refers to steps that may be taken deliberately to increase the effectiveness and the rate at which an individual assimilates new knowledge or new techniques. Experience refers to the cumulative effect of exposure to events or occurrences without any systematic attempt on the part of another person to focus attention upon what is most important."

The effectiveness of training to increase the consistency of assessments by the laboratory panel was demonstrated by Ehrenberg and Shewan (17) in experimental work with fish.

Hanson, Davis, et al (23) found that ratings given foods in consumer-type tests are influenced by previously tested samples, even though samples are presented singly

up to a two-day interval between tests. There was a greater tendency for tasters to rate succeeding samples lower rather than higher than the first sample. It is suggested that this may be the result of development of a more critical attitude on the part of judges with continued testing. These findings are similar to those of Dove (16) who previously had shown that ratings given a food are influenced by the quality of other samples presented simultaneously. Peryam and Girardot (50), in describing the use of the hedonic scale, recognized that samples presented together may mutually influence the ratings obtained for each by the tasters.

The effect of age of persons was tentatively evaluated by Polemis (54) in connection with the findings of five flavor preference attitude surveys conducted with army personnel in 1950 and 1951. These surveys were based on the use of questionnaires and did not involve actual product tasting. Her reports were to the effect that level of preferences was generally highest for the youngest soldiers but varied with specific foods, such as, acceptance in the case of vegetables tended to be greater among the older groups, whereas acceptance for desserts was greater among the younger groups. These results are in general agreement with the findings of Abbott, et al (1), and Eppright (19).

Area from which the soldiers came was also examined relative to likes and dislikes for certain kinds of food in this same study. The results strongly suggested that likes or dislikes for certain kinds of food may be related to the region from which a soldier originated. For instance, roast lamb seemed to be liked very much by soldiers from New England but liked to a lesser extent by soldiers from the Northwest.

Baten (2) made a comparison between the ability of college freshmen, average age of 18, and college faculty members and secretaries, ages 35 to 40, to differentiate levels of salt-treated tomato juice, levels of sugar-treated apple juice, and concentrations of lemonade. It was found that the younger tasters were more discriminating of saltiness and sweetness differences than the older group, but that discrimination of sourness was equal in both of the groups. The younger tasters showed less variation in their evaluations than the older tasters. It was somewhat cautiously concluded that younger and more mature tasters probably do not evaluate certain foods with the same degree of discrimination and that young people may distinguish some flavor differences when mature people do not.

Laird and Breen (36) examined the preferences of three age groups with respect to pineapple juice of five

different degrees of sweetness. From their work they concluded that preferences for sweetness and tartness among sixty boys and girls, ages 12 to 18, paralleled preferences among fifty men and women, ages 20 to 40. Most choices in both groups were for sweeter pineapple juice. Among fifty men and women, ages 50 to 68, however, they found that preferences for sweetness decreased and preferences for tartness increased in comparison with the younger groups.

Suggestion, Knowledge, and Related Factors

Foster, Pratt and Schwartz (20) investigated the influence of the voiced responses of taste panel members on other panel members in a comparison of the "round-table" test situation with the individual testing station method. The individual testing method isolates the taster from the other members of the test panel, whereas in the round-table method reports are orally given and decisions arrived at by discussion. The work of these investigators showed that the responses of observers in a round-table situation can be influenced by the report of the first observer. The biasing influence was not apparent with every product tested, but it was demonstrated that qualitative flavor differences and preferences can be created where none exist, that is, in samples which are identical. They concluded that round-table testing methods may result

in data of only precarious reliability, because of the potential influence of the report of a single member on the entire group, although they pointed out that such round-table environment may most closely approach the home consumer situation.

Harper (24) suggests that the following factors may be of importance in human judgments:

(a) The influence of "irrelevant characteristics," referring to the effect that personal judgment of one characteristic may have to influence the judgments by the same person for other characteristics, sometimes known as the "halo effect" by psychologists;

(b) The apparent tendency of judges to avoid the use of extreme categories unless care is taken in the organization of adequate judgment scales;

(c) The effects of "serial events" on judgment, such as the influence of judgments made in the recent past on present judgments, accentuation of "contrast" between samples when they are in close proximity (such as, the exaggeration of sourness after tasting something sweet), over-estimations and under-estimations in ascending and descending series;

(d) The potential influence of sample labels on judgments, signifying the need, as suggested by Hening (26), for careful use of non-systematic letters and numbers

in identifying the items for judgment;

(e) The influence of knowledge of results on improving performance in certain judgment situations;

(f) The influence of expectation and foreknowledge on human judgment, as in the instance of judgments likely being "easier and more correct if the person making them knows the general type of variation which he is about to meet."

Josephson (32) maintains that "the average human being is equipped with an extremely sensitive taste mechanism, with which to evaluate flavor, as well as a taste memory to which he can refer in establishing his acceptance of a food." He describes the taste mechanism as consisting of gustatory, tactual and olfactory senses which depend upon chemical stimuli for the composite sensory response described as flavor. The individual does not, however, voluntarily sort out the various flavor qualities prompted by the taste mechanisms, but involuntarily evaluates them as an over-all response, influenced whenever such exists for any food by an established taste memory.

Crocker, Sjostrom and Tallman (13) believe that taster's knowledge, or guess, as to the source or maker of the product may cause assumptions to be made concerning flavor and quality which will influence his opinion. The physical setting of presenting a product, the questions

asked, the personality and attitude of the investigator may, as they see it, all recall past experiences and influence the consumer's interpretation of his immediate flavor experience.

Various means of identifying or coding samples which are presented to tasters are commonly used in flavor testing. Single or multiple letters, numbers, or combinations of letters and numbers may be used. Color designations or geometrical shapes or patterns are preferred by some workers. Published experimental evidences on the actual effect of using various coding systems in flavor testing are few, with choices of a particular system apparently based largely upon logic and experience. Ishler, Laue and Janisch (30) state that in their contacts with commercial survey organizations it is reported by some that consumers exhibit bias in preference toward familiar letters and toward even numbers. Others observe that biases exist for the letters at the beginning and at the end of the alphabet and that numbers can introduce bias on the basis of their order implication. These workers conducted a series of store tests and home use tests with a soluble beverage to see what could be learned concerning biasing effects of individual letter codes. They found that code letter bias did not appear in the store tests, which they attributed to the particular conditions of their store

test. Bias was indicated, however, toward certain single letters in the home use tests. They suggested means of minimizing this bias by either reversing codes between halves of the same test, or by using only codes for which no appreciable bias has been found.

Physiology, Sensitivity, Sex

Beidler (3) discusses physiological factors involved in taste perception, relating them principally to the structure of the taste organs, nerve pathways involved and possible chemical reactions connected with salty, sour, bitter and sweet taste sensations. He points out variations in the ability of various animals to respond to certain chemical substances. The striking segregation of people into those to whom the chemical substance phenylthiocarbamate elicits a bitter response and those to whom it elicits a tasteless response is described, emphasizing physiological variation in taste. This lack of "bitter" sensitivity is believed to be inherited.

Morgan (47) suggests that structurally taste senses should probably be grouped with skin senses rather than with smell. Taste "buds" are clusters of cells embedded in ridges of the tongue and from which nerve fibers lead to branches of three cranial nerves. Gustatory sensitivity is a function of the penetration of materials into

the taste cell, entering through a pore in the taste bud, and of the rate of consequent breakdown of chemical substances in the taste cell. It is strongly believed that penetrability is an important factor in the development of taste sensation and in the effectiveness of a taste stimulus.

Qualities of taste have been found by Pfaffman (52) to be closely related to three types of nerve fibers which are responsive to different kinds of stimuli. Working with the taste nerves of the cat, he found that certain nerves gave impulses only when acid was presented, that other nerves responded when acid or salt was the stimulus, and that still other nerves yielded impulses to acid or quinine. This work lends strong inferential support to the common concept that human taste falls into four primary groups--sweet, sour, salty, and bitter, although it points out that it may be in error to assume that there is one kind of taste receptor for each taste quality perceived. Rather, there may be several receptors acting together in a particular pattern to produce a recognizable taste sensation.

Woodworth (60) reports the work of Mayer, in 1927, to show that adaptation to taste sensations may occur among tasters. In an experiment with three trained individuals, thresholds were found to rise tenfold to the recognition

of a sweet (sugar) solution after holding a cane sugar solution in the mouth for two minutes. This meant that it required ten times as strong a concentration of sugar solution to invoke a sweet taste subsequent to holding the adapting solution in the mouth. Adaptation increased directly with the concentration of the adapting sugar solution or with the time of holding in the mouth up to two minutes. A bitter solution (quinine) produced lesser modification of taste sensitivity and common salt and hydrochloric acid resulted in only slight adaptation. Recovery to normal from adaptation was found to occur gradually over a ten to fifteen minute period.

The pleasantness or unpleasantness of the four primary taste sensations was examined by Engel (18). He presented solutions of varying concentrations of cane sugar, tartaric acid, common salt, and quinine to fifteen individuals for their judgment as pleasant, unpleasant, or neutral. He concluded from the results that each sensation increased slightly in pleasantness up to a point at which the solution was of a certain strength typical for each, following which sour, salty, and bitter fell off rapidly to increasing unpleasantness as the concentration increased. Sweet on the other hand did not ever reach unpleasantness except after satiety to a few individuals, maintaining generally a high degree of pleasantness, which reached a

maximum at about nine per cent solution and diminished gradually, but only slightly, thereafter.

Mackey and Jones (40) did not find a correlation between sensitivity of tasters measured against water solutions of primary taste substances and the ability of the same tasters to arrange in order foods (applesauce, pumpkin, and mayonnaise) containing various concentrations of these same taste substances. King (33) reported similar findings in testing a group of tasters for sensitivity to pure solutions of sodium chloride, lactic acid, sucrose and caffeine, followed by discrimination tests on bread containing different levels of these same taste substances. This investigator concluded that there was no correlation between the consistency of the group in judging pure solutions and its ability to rate correctly bread samples containing various quantities of these substances.

Tasting performance of men in contrast to women has been scrutinized but apparently without decisive evidence of important distinctions between them in food preferences. Hopkins (29) reported that women had significantly lower taste sensitivity than men. Weckel (59) reported that men excelled in identifying solutions by taste while women excelled in identifying odors. Bliss, et al (6) found that in judging dairy products, differences in flavor preferences between the sexes were not large enough to be

considered. Langwill (37) observed in a consumer preference test that more than half of both men and women preferred moderately sweet and salty foods, that more women than men preferred highly salty and sour foods, that more than half of the men liked slightly sour foods, and that women showed greater sensitivity of taste in distinguishing between sweet, salty, sour, and bitter. Laird and Breen (36) indicated that the women of three age groups as compared to the men appeared to have more preference for tart and less for sweet, although this was more pronounced in the older age group. Differences in the number and functioning of taste buds have not been established between sexes, according to their report.

Rinsing, Swallowing, Fatigue, Hunger

Studies which have been reported on the effects of rinsing and swallowing have dealt primarily with flavor difference testing. Some investigators have reported that it is better for taste judges to eject samples rather than to swallow them (9, 13, 45, 53). Lukewarm water is recommended for rinsing the mouth between samples (39, 45, 53) and has been so used in taste testing various products. Wiping the tongue with a napkin followed by a water rinse was suggested for egg products (21, 41). Crocker (9) believes that something may be gained by the use of

various artifices between tasting samples, particularly rinsing with water, to remove adherent material, but that generally the best technique is to allow the saliva to lave the taste buds in a natural manner. He expresses the opinion that rinsing of the mouth usually is not necessary or desirable, except after tasting clinging substances. Trout and Sharp (56) suggest that tasting a poor-flavor sample without having previously rinsed the mouth may produce discrepancy in results. Other workers have considered that rinsing is not particularly helpful in their taste tests (46).

Fatigue in tasting is considered a cause for concern in certain types of taste testing. It appears to be characterized by a diminishing ability of the taster to recognize small differences in flavor when tasting is prolonged. Taste panel members sometimes complain of a sensation of fatigue. Bengtsson and Holm (4) report that fatigue is due to reduced functioning of the gustatory and olfactory nerves. They suggest that the stronger the taste and odor, the fewer samples should be tasted, and that the number of samples presented may be increased if tastes or odors are not pronounced. Clendenning (8) attributes fatigue to exhaustion of sensory end organs when stimulation is repeated. He recommends that as few samples as possible be presented to the taster. Peryam

and Swartz (51) indicate that if samples contain flavor characteristics of low impact or intensity, one triangular test may be followed by a second if a short rest is taken between them. Crocker (9) says that recovery to normal sensitivity after stimulation is a matter of minutes. He suggests that tasting be slow, with rest periods of varying duration between tastes, depending on the material.

Lane, Ishler and Bullman (38) tested the effects of fatigue in double triangular tests of a soluble beverage and of maple syrup. They found no evidence of fatigue in the beverage tests, but observed diminishing success in the second tests of the maple syrup samples. They concluded that fatigue may be a prominent factor in tasting of some types of food but probably of no importance in other foods. It is suggested that specific investigation is needed to establish the role of fatigue with regard to specific products.

Meyer (43) investigated hunger as a factor in taste sensitivity of humans. Using nine volunteer students, he ascertained that thresholds of sensitivity to solutions of sodium chloride (salty), quinine sulfate (bitter), and glucose (sweet) remained constant over a two-day period of abstinence from food.

Similar conclusions were reached by Janowitz and Grossman (31) in examining the recognition thresholds of

small groups of medical students at the University of Illinois. Based upon that point at which a substance was recognized, they determined that the threshold for a salty and a sweet solution and for the aroma of coffee did not differ before and after eating or before and after suppression of the appetite by a chemical substance. Their work showed that there were clear differences in the taste and smell thresholds of individual tasters.

Color, Appearance, Sound

It is commonly considered that color is an important aspect of the acceptance of foods. Reports of research as to the nature of the influence of color are, however, sparse. General recommendations in the conduct of flavor difference tests are usually to eliminate the factor of color differences by using special lights, dyeing the samples, or training the judges to taste without looking at the sample (14, p.85). Honing (26) recommends that in some flavor testing tasting should be done in a darkened room to cover up differences in color of the samples. Clendenning (8) and Hicks (27) recommend that controlled lighting is desirable and that brilliant lights may be detrimental to best taste test results.

Kanig (42) reports that green seems to be the most agreeable of colors for foods. Blue and violet tend to be

in the "inedible" class, with shades of these colors more acceptable in foods than their pure hues. "Blue," he states, "is the real gastronomic taboo."

He found color preferences to vary according to localities. Bostonians purchased brown eggs while New Yorkers bought white eggs. Deep-yellow butter is more popular in the midwest than in the east. Solid-yellow scrambled eggs find acceptance in San Francisco while pale scrambled eggs are desired in New York.

He suggests that food-color associations are deep-seated and must be dealt with carefully. Tomato-colored bread was tried on a midwestern market only to find that people desired to buy "bread that looked like bread."

The effects of color, as well as consistency, to influence flavor are suggested by Crocker, Sjostrom and Tallman (13), who mention that appearance may predispose an invalid judgment, and that consumers are prone to relate flavor type or strength with the character or intensity of color. Consistency, or apparent consistency, of a product, they feel, may introduce psychological implications and invoke prejudice in flavor judgment.

Crocker (11, 12) states that the three senses--taste, smell, and feel--are the principal ones involved in the flavor of food, but that we taste first with our eyes. "The food," he says, "has to look right, to look as though

it is clean and fit to eat. Having assured ourselves of the appearance, then we appraise by the two chemical senses, taste and smell, and to some extent by feeling."

He contends that sound has influence too; that a loud noise may prevent entirely our ability to smell or taste, yet "softly played dinner music can create an environment favorable for elegant dining." He suggests that the act of eating may itself produce characteristically important sounds, such as the crunching of an apple or the snap of crackers or "even the silence that goes with indulgence in whipped cream."

EXPERIMENTAL APPROACH

Objectives

The aim of this research was to ascertain the influence of certain psychological and environmental factors on the results obtained in flavor panel preference testing. Determination of this influence was considered important in order to increase confidence in this type of flavor testing and to aid in the development of improved techniques. The factors chosen for investigation were those which appeared to the writer to be the least explored, or to be surrounded by confusion or controversy in literature reports, and for which available facilities could, it was believed, provide suitable tools for investigation. The valid conduct of flavor preference tests could be strengthened by a more thorough understanding of the effect of factors involved in flavor preference.

Seven factors were chosen for examination. These were: (1) The number of samples presented simultaneously; (2) The location of the test; (3) Sound accompanying the test; (4) Information possessed by the taster; (5) Rinsing the mouth with water between tastes; (6) Swallowing the product; (7) The amount of sample consumed by the taster.

Method of Evaluation

The method used for conducting all of the experiments in this research was the reference-type, multiple-sample preference test, utilizing a two-directional, seven point word-phrase scale. In this method samples of food in small containers are served on a tray to the tasters. One of the samples presented is identified to the taster as the reference product with which he is to compare each of the accompanying samples individually and directly. On an appropriate ballot provided the taster records his judgment as to whether he considers the specific sample slightly, moderately, or much better or poorer, or neither better nor poorer, than the reference. In essence, this is the paired preference test with an identified reference.

Scoring System

The judgments which the tasters indicate on the ballots are transposed to numerical scores preliminary to analysis. These scores are arrived at on the basis of a seven point scale, in which 1 represents "much poorer flavor than reference," 2 represents "moderately poorer flavor," 3 represents "slightly poorer flavor," 4 represents "neither better nor poorer flavor," 5 represents "slightly better flavor," 6 represents "moderately better

flavor," and 7 represents "much better flavor than reference."

Flavor Panels

All panels of tasters for these tests were composed of men and women Oregon State College students. They were not pre-selected for the tests nor were they screened nor trained. They participated solely on the basis of voluntarily responding to a generalized call for tasters. They were paid a small amount for their time and were rewarded on occasions by treats of candies or refreshments. Their only other incentive was interest. The effort was made at all times to have each panel as large as possible. The limiting factor was generally the availability of tasters under the times and conditions of a specific test.

Product

The selection of a food product for the investigation of the chosen psychological and environmental factors involved in flavor preference testing was based on the premise that such a product should be as homogeneous as possible under conditions of serving, should lend itself to quantity kitchen preparation, should be relatively stable, and should be generally familiar to most people. A series of small informal taste tests were conducted with

a small group of people on products including orangeade, orange soda, vegetable juice, tomato juice, apricot juice, pear juice, apple juice, and prune juice. It will be noted that these are all beverages, whose homogeneous nature seemed by logic to fulfil the first criterion for a suitable product. The outcome of these informal evaluations was to convince the writer that tomato juice was a very suitable medium for the conduct of these experiments. Tomato juice possessed a valuable feature, in addition to the other characteristics sought, in that it could be obtained on the market in varying grades of quality. Of course, the selection of this single product for the conduct of this research necessarily implies that conclusions reached from it may have full force only with tomato juice. But it was believed that a greater breadth of information regarding psychological and environmental factors could be obtained by concentrating on one product. The use of more than one product would make the experimentation too bulky to be practical and appeared unnecessary to achieve the goals of this inquiry.

THE EFFECT ON FLAVOR PREFERENCE OF THE NUMBER OF SAMPLES PRESENTED SIMULTANEOUSLY

Procedure

For the purpose of ascertaining whether two, three, or four samples, plus a reference, presented at one time on a single tray would affect flavor panel preference results, fourteen trays of tomato juice samples were served to each taster of a panel of eighty-four men and women college students. These were presented on seven different days, two trays per day. Each tray of the fourteen contained a reference sample, identified by R marked on the cup, a coded cup containing identical juice to the reference sample, and one, two, or three additional coded cups which contained tomato juice of brands different from the reference and the coded duplicate. This provided variations in the number of samples presented of two samples plus reference, three samples plus reference, and four samples plus reference.

The reference and the coded duplicates consisted of brand K tomato juice while the additional coded cups contained juice of brands A, B or C. The seven combinations of KA, KB, KC, KAB, KAC, KBC and KABC were served twice during the test series. The order of presentation was completely randomized. The use of the different brands,

which were selected on the basis of preliminary tests for their preference values, provided variations in the kind of samples accompanying brand K in the two-sample and three-sample groups.

The tomato juice for these tests was obtained from single lots of commercially packed juice. On each testing day throughout the series the contents of a uniform number of cans from one lot were blended together for a single brand, maintained at a uniform room temperature of seventy degrees Fahrenheit, plus or minus three degrees, and served from a covered pan. Two ounces of juice were placed in the reference cup, while one ounce of juice was placed in the coded cups. The code numbers, of four digits, were obtained from a random number table.

The tasters were seated in individual taste testing booths which were lighted uniformly by green incandescent light. The green light masked any possible differences in color of the samples. The trays were passed to the tasters through a sliding-door opening, which was closed during the tasting. Each booth was equipped with a pencil, a service counter, a stool, a porcelain basin containing hot and cold water supply, and paper rinsing cups for use by the tasters at their own discretion.

A printed ballot was given to each taster with each tray. This ballot, shown in Exhibit A, instructed the

Exhibit A

-- Flavor Test Ballot --

Date:_____ Product:_____ Name:_____

Test Number:_____

INSTRUCTIONS:

The side dish contains a reference sample. Compare the flavor of each sample directly with the flavor of the reference sample.

(a) Taste the reference and the sample as many times as necessary;

(b) Write the number which is on sample container in a space after the statement which indicates your judgment:

Much better flavor than reference _____

Moderately better flavor _____

Slightly better flavor _____

Neither better nor poorer flavor _____

Slightly poorer flavor _____

Moderately poorer flavor _____

Much poorer flavor than reference _____

Comments:

Be sure your name is written on this ballot. When finished, please open the sliding door, push tray through, fold the ballot once and place it in ballot box.

THANK YOU.

taster to compare the flavor of each coded sample directly with the flavor of the reference sample, then to record his judgment, in columns provided, concerning whether he considered the flavor of the coded samples slightly, moderately, or much better or poorer than the flavor of the reference sample, or neither better nor poorer than the flavor of the reference.

Analyses and Results

The scores for the coded duplicate samples were tabulated on the basis of the seven variations under which they were obtained. Inasmuch as these samples were identical to the reference, it would be expected that their scores would occupy, within the range of chance variation, the center position on the scale signifying "neither better nor poorer flavor." These scores were statistically analyzed for variance.

The analysis of variance, seen as Table 1, indicated that significant differences existed among the mean scores obtained for the duplicate samples when presented with different number of samples. There was, however, no significant difference in the mean scores obtained in the presence of differing brands of accompanying samples. The general mean score also differed significantly from the value 4, expected as the true score for the preference

between identical samples.

Table 1

Effect of Number of Samples Presented Simultaneously

Variation Source	Degree of Freedom	Mean Square	F
Deviation from score 4	1	12.8648	17.06**
KA, KB, KC	2	0.0774	0.10
KAB, KAC, KBC	2	0.6607	0.88
Number of Samples	2	5.8266	7.72**
Tasters	83	2.7237	3.61**
Combinations x Tasters	498	0.9587	1.27
Error	588	0.7543	----
Total	1176	----	

**Significant at 1%

The mean scores obtained under conditions of two-sample, three-sample, and four-sample presentation, seen in Table 2, were also analyzed for their deviation from the true value of 4. This revealed that the mean score in the case of the three-sample plus reference combinations was clearly significantly different from the true value.

The mean preference scores for the accompanying brands of tomato juice, which were found to have no effect on the findings of the panel for the duplicate of the reference, were as follows: Brand A, 3.86; Brand B, 3.61; Brand C, 1.40.

Table 2

Effect of Number of Samples
Mean Scores of Duplicate Samples to the Reference

Brands	Mean Score	Number of Observations
KA	4.01	168
KB	4.02	168
KC	3.98	168
Two-sample Combination:	4.01	504
KAB	3.79	168
KAC	3.86	168
KBC	3.73	168
Three-sample Combination:	3.79**	504
KABC	3.88	168

**Significantly different from 4 at 1%

Summary

It is seen from the analyses that the mean score of this panel was influenced by the number of samples presented simultaneously on a single tray. At the same time, the scores were not influenced by the accompanying samples which were presented.

Only in the case of two samples plus a reference was the panel preference judgment clearly equivalent to the true value. This was the smallest number of samples presented simultaneously on one tray in this experiment, which would imply that greater number of simultaneous samples tends to lessen the accuracy of judgment.

In the case of three samples plus a reference, it was clear that the preference finding of this panel was different from the known true value. With respect to four samples plus a reference, the score of the panel was not as clearly different from the true value. This must be viewed in the light of the number of observations on which the judgment was based. The four-sample panel judgment was based on 168 observations, whereas the three-sample and two-sample judgments were based on 504 observations each. This reduces the comparative dependability of the four-sample score.

The magnitude of differences from true value is noted to be relatively small--0.21 in the case of three samples plus reference. This may not have serious practical implications on the seven point scale used, in which 1.00 represents the difference between "neither better nor poorer flavor" and "slightly" better or poorer. Nevertheless, the difference is real on the basis of the analysis, and it is clear that greatest preference accuracy was obtained in the case of the fewest number of samples presented simultaneously.

THE EFFECT ON FLAVOR PREFERENCE OF THE LOCATION OF THE TEST AND OF SOUND ACCOMPANYING THE TEST

Procedure

The effects of location and sound on results obtained by the flavor preference panel were investigated by conducting a series of six flavor tests, on six days, in three different locations under conditions of sound and quiet. The same tasters were used throughout the series. By presenting two different trays of samples at each sitting, two sets of observations were obtained for evaluation.

At each test, three cups of tomato juice were served on a single tray to each taster. One of the cups, marked R and containing two ounces of juice, was identified as the reference sample. The other two cups, each containing one and one-half ounces of juice, were marked only with four-digit code numbers, selected from a random number table.

The tomato juice for these tests was obtained from single lots of three brands of commercially packed juice. On one tray of the two trays presented at a sitting, the reference and one of the coded samples consisted of brand A. The other coded sample was brand B. On the other tray, the reference and one coded sample were brand C, while the

other coded sample was brand B. The order in which the two trays were presented was randomized from taster to taster.

On each testing day, the juice was prepared for serving by blending the contents of a uniform number of forty-six ounce cans taken from the same cases throughout the series of the same lot (identified by packer's code number) for brand A, brand B, and brand C, separately. The juice was held during each test in the large blending pans and served at room temperature of seventy degrees Fahrenheit, plus or minus three degrees. The pans were kept covered during the tests except when the product was being dispensed by ladle from them.

A printed ballot was given to each taster with each tray of samples (Exhibit A). This ballot instructed the taster to compare the flavor of each coded sample directly with the flavor of the reference sample, then to record in the columns provided his judgment as to whether he considered the flavor slightly, moderately, or much better or poorer than the reference flavor, or neither better nor poorer than the flavor of the reference.

Locations

Three locations were chosen for the tests in proximity to the preparation kitchens of the food technology

building, Oregon State College. These consisted of: (a) Individual taste testing booths equipped with an individual basin, stool, service counter, and sliding-door opening through which the samples were served; (b) An open rectangular room approximately twenty-five feet wide and thirty feet long, surrounded on three walls with an open counter on which the samples were served to the tasters as they were seated on stools; (c) An auditorium approximately seventy-five feet long, fifty feet wide, and twenty-five feet high, at the front of which three rectangular tables were placed on which the samples were served to the seated tasters.

Each location was lighted for the test by white incandescent room lights. Room temperatures were approximately the same in each due to a central circulating air system in the building. Printed placards were placed prominently on the walls instructing the tasters not to converse during the tests. No verbal instructions were given to the tasters by the serving attendants.

Sound

Each of the locations was used on two days. On one of these days quiet was maintained throughout the entire test. Prominent placards on the walls instructed tasters and serving attendants to maintain silence. On the other

day, controlled sound was introduced into the tasting area. Placards requesting no conversation remained on the wall at this time.

The controlled sound consisted of a tape recording of the noises of a large restaurant. This was played continuously throughout the test on an Eicor, Model 230, tape recorder, which was hidden from the view of the tasters. The tape recording was prepared from a commercial transcription with the aid of the technical staff of KOAC radio station, Oregon State College, Corvallis. Sounds introduced by it were the intermixture of clattering dishes and utensils, scuffing and scraping, voices, and background noises. No intelligible words were heard. Intensity of the sound in each location was maintained at a uniform level of eighty decibels, plus or minus five decibels, established by means of meter measurement with respect to the positions occupied by the tasters while tasting.

The order in which the six tests comprising the three locations under conditions of sound and quiet were conducted was randomized. Eighty-four tasters rendered judgments under the six conditions for products CCB, while eighty-three judgments were obtained under the six conditions of location and sound for the products AAB.

Analyses and Results

The scores obtained for all coded samples were tabulated according to the conditions of location and sound. Scores for the coded samples which were identical to the reference were established on the basis of the seven points of the ballot scale. Scores for the coded samples which were not identical to the reference were calculated on the basis of their difference from the score for the identical sample. This provided an anchor for these judgments, inasmuch as their true value could not be known. Effects of the testing conditions on the preference for these samples could be viewed in terms of whether the difference in preference between them and the coded control samples was caused to change.

Analyses of variance were conducted for the scores of the samples which were identical to the reference and for the scores of the non-identical samples in each tray combination. The analysis for the scores of the coded identical samples in both combinations revealed no significant differences in the panel preference judgment obtained under the various conditions of location and sound or quiet. The analyses, shown in Tables 3 and 4, indicated that there were some differences in preferences among individual tasters and that in the case of product C there were significant interactions between the tasters

and the locations, the tasters and the sound, and the sound and the locations, implying that the reactions of tasters were not all the same under the conditions. Nevertheless, the preference findings of the panel remained the same in the three locations and whether there was sound or quiet.

Table 3

Effect of Location and Sound
Duplicate Sample A

Variation Source	Degree of Freedom	Mean Square	F
Sound	1	0.3936	0.44
Location	2	0.6888	0.77
Sound x Location	2	1.5201	1.69
Tasters	82	2.1503	2.39**
Sound x Tasters	82	0.9261	1.03
Locations x Tasters	164	0.6786	0.75
Error	164	0.9002	----
Total	497	----	

**Significant at 1%

Table 4

Effect of Location and Sound
Duplicate Sample C

Variation Source	Degree of Freedom	Mean Square	F
Sound	1	0.8750	1.81
Location	2	1.0139	2.10
Sound x Location	2	1.6250	3.37*
Tasters	83	1.1380	2.36**
Sound x Tasters	83	2.2083	4.58**
Location x Tasters	166	1.7267	3.58**
Error	166	0.4824	----
Total	503	----	

*Significant at 5%. **Significant at 1%.

Analysis of variance of the scores of coded sample B in both combinations also indicated no significant differences in mean scores of the panel obtained under the various locations or conditions of sound or quiet. The analyses, shown in Tables 5 and 6, again revealed that all tasters did not agree with each other in their preferences. The general mean scores for product B were 5.57 with reference A and 5.84 with reference C.

Table 5

Effect of Location and Sound
Sample B with Reference A

Variation Source	Degree of Freedom	Mean Square	F
Sound	1	0.0984	0.06
Location	2	0.2892	0.17
Sound x Location	2	4.9719	2.85
Tasters	82	3.7366	2.14**
Sound x Tasters	82	1.5293	0.88
Location x Tasters	164	1.7160	0.98
Error	164	1.7442	----
Total	497	----	

**Significant at 1%

Table 6

Effect of Location and Sound
Sample B with Reference C

Variation Source	Degree of Freedom	Mean Square	F
Sound	1	0.6429	0.31
Location	2	5.9068	2.84
Sound x Location	2	1.1486	0.55
Tasters	83	5.5067	2.65**
Sound x Tasters	83	2.2041	1.06
Location x Tasters	166	2.0207	0.97
Error	166	2.0791	----
Total	503	----	

**Significant at 1%

Summary

These tests demonstrate that the three locations in which the flavor tests were conducted did not alter the preference findings obtained from the taste panel, and that the presence of generalized sound which is typical of a large restaurant, as contrasted to quiet, did not affect the preference results obtained. The analyses indicate that all tasters did not agree in their individual preferences, but panel preference did not change from location to location nor with sound or quiet. Flavor preference for tomato juice by this panel could be as well ascertained in any of the three locations examined and without concern for restaurant-like sound or maximum quiet.

THE EFFECT ON FLAVOR PREFERENCE OF INFORMATION POSSESSED BY THE TASTER

Procedure

The influence of information on flavor panel preference was examined by conducting seven separate flavor tests using three different panels of tasters on three days. Each test of the series consisted of serving simultaneously on a tray to each taster three portions of identical tomato juice, accompanied by specific printed information concerning certain of them. One of these three samples was identified as reference by R marked on the cup. The other two samples were identified by four-digit code numbers, obtained from a random number table. Two ounces of tomato juice were contained in the reference cup, while the coded cups each contained one and one-half ounces of tomato juice. The tasters were seated in individual tasting booths which were lighted uniformly with white incandescent light and equipped with individual porcelain sinks, hot and cold water supply, pencils, service counter, and paper rinsing cups to be used at the taster's own discretion.

Two printed ballots were given to each taster with each tray of samples. One of the ballots was identified with one of the coded cups on the tray and the other

ballot was identified with the other coded cup. In addition to giving uniform instructions concerning the procedure of tasting and the use of scoring columns, each ballot contained a specific statement of information related to the code number of the cup with which it was identified. The order in which the two ballots were presented to each taster was randomized.

The general instructions on the ballot (Exhibit B) directed the taster to compare the flavor of the sample contained in a specifically numbered cup with the flavor of the sample in the reference cup, then to indicate in the column provided whether the flavor of the coded sample was slightly, moderately, or much better or poorer, or neither better nor poorer, than the flavor of the reference sample.

In a prominent position immediately adjacent to the scoring column, a statement was printed conveying specific information concerning the numbered sample being tasted. These statements for each pair of coded samples presented on a single tray were chosen to be contrasting in nature. It was intended that they provide the taster with differing information about some aspect of raw product quality or processing treatment in connection with the two coded (identical) samples being flavor tested at one time.

Exhibit B

-- Flavor Test Ballot --

Date:_____ Product:_____ Name:_____

INSTRUCTIONS:

The cup labeled "R" contains a reference sample. Compare the flavor of sample_____ directly with the flavor of the reference sample.

(a) Taste the reference and the sample as many times as necessary.

(b) Check the space on this ballot which indicates your judgment concerning this sample:

Much better flavor than reference _____

Moderately better flavor _____ ***

Slightly better flavor _____ (Specific Informational statement is written here concerning sample numbered above)

Neither better nor poorer flavor _____

Slightly poorer flavor _____

Moderately poorer flavor _____ ***

Much poorer flavor than reference _____

Comments:

When finished, please open sliding door, push cups through. Place ballot in ballot box. Thank you.

BE SURE YOUR NAME IS WRITTEN ON THIS BALLOT

The statements of information formulated to provide pairs of contrasting information were the following:

Test 1: (a) Sample 3586 consists of tomato juice prepared from CAREFULLY SELECTED, SUNRIPENED, FANCY GRADE tomatoes.
(b) Sample 4296 consists of tomato juice prepared from UNSORTED, RAIN-DAMAGED, SUB-STANDARD tomatoes.

Test 2: (a) Sample 0715 consists of CHEMICALLY TREATED tomato juice.
(b) Sample 9303 consists of tomato juice which has NOT BEEN TREATED with chemicals.

Test 3: (a) Sample 6850 consists of tomato juice which HAS BEEN EXPOSED TO RAYS for the purpose of sterilization.
(b) Sample 6147 consists of tomato juice which has NOT BEEN EXPOSED TO RAYS for sterilization.

Test 4: (a) Sample 4235 consists of tomato juice which has had MONOSODIUM GLUTAMATE added to it.
(b) Sample 5693 consists of tomato juice which has not had Monosodium Glutamate added to it.

Test 5: (a) Sample 7926 consists of tomato juice which has had SALT added to it.
(b) Sample 4519 consists of tomato juice which has not had Salt added to it.

Test 6: (a) Sample 2736 consists of tomato juice which has had SUGAR added to it.
(b) Sample 5906 consists of tomato juice which has not had Sugar added to it.

Test 7: (a) Sample 9164 consists of tomato juice which has had LEMON JUICE added to it.
(b) Sample 8360 consists of tomato juice which has not had Lemon Juice added to it.

The first three tests were conducted at a single sitting with a taste panel of 132 college students. Tests 4 and 5 were conducted at a single sitting with a panel of 49 college students. Tests 6 and 7 were conducted at a

single sitting with a taste panel of 39 college students. The order in which the test trays, containing samples and paired ballots, were presented in each instance was randomized from taster to taster. No information was given to the tasters pertaining to the nature of the tests other than appeared on the printed ballots. No verbal instructions were delivered by the serving attendants.

Analyses and Results

The scores obtained from the taste panels for the coded samples were tabulated and analyzed for each test separately by statistical analysis of variance.

The analyses of variance of these scores revealed the following effects of the ballot information on the preference findings:

1. Information concerning the quality of the raw product clearly influenced the panel preference (Table 7).

2. Information concerning treatment by unspecified chemicals did not affect the panel preference (Table 8).

3. Information concerning exposure to unspecified rays for sterilization had no influence on the panel preference (Table 9).

4. Information concerning the addition of monosodium glutamate had no effect on the panel preference (Table 10).

5. Information concerning the addition of salt to

the tomato juice clearly influenced the panel preference (Table 11).

6. Information that sugar was added to the tomato juice changed the panel preference (Table 12).

7. Information that lemon juice was added to the juice altered the panel preference (Table 13).

The mean scores, seen in Table 14, indicate that the information given produced the following changes:

(a) Information concerning high quality raw product raised the panel preference, whereas information of low qaulity raw product lowered the preference.

(b) Information that salt was added raised the preference, while the statement that salt had not been added lowered the preference.

(c) Information that sugar was added raised the panel preference, and the statement that sugar had not been added lowered the preference.

(d) Information that lemon juice was added raised the preference, while the statement that lemon juice had not been added lowered the panel preference.

The analyses indicate also that the scores of some of the individual tasters significantly differed with respect to the tests involving information of treatment by chemicals, exposure to rays, and quality of raw product, but that the scores of individual tasters did not

significantly differ in the tests involving information regarding the addition of lemon juice, sugar, salt, and monosodium glutamate. This would signify agreement among the tasters with respect to either the effect or the lack of effect of the information on lemon juice, sugar, salt, and monosodium glutamate, strengthening the conclusions reached concerning them. It would suggest, however, that in the cases of chemicals, rays, and raw product quality, it is possible that some tasters may have been influenced by the information while others were not. Although there was lack of agreement in preference among some tasters in the case of the raw product information, the flavor panel preference score was clearly altered by the information.

Table 7

Effect of Information re: Raw Product Quality

Variation Source	Degree of Freedom	Mean Square	F
Quality Information	1	28.6705	28.93**
Tasters	131	1.3591	1.37*
Error	131	0.9911	----
Total	263		

*Significant at 5%. **Significant at 1%.

Table 8

Effect of Information re: Chemical Treatment

Variation Source	Degree of Freedom	Mean Square	F
Chemical Information	1	0.0000	0.00
Tasters	131	1.5562	1.32*
Error	131	1.1756	----
Total	263	----	

*Significant at 5%

Table 9

Effect of Information re: Exposure to Rays

Variation Source	Degree of Freedom	Mean Square	F
Rays Information	1	4.3788	3.78
Tasters	131	1.6775	1.45*
Error	131	1.1574	----
Total	263	----	

*Significant at 5%

Table 10

Effect of Information re: Addition of MSG

Variation Source	Degree of Freedom	Mean Square	F
MSG Information	1	0.3674	0.28
Tasters	48	1.8525	1.42
Error	48	1.3048	----
Total	97	----	

Table 11

Effect of Information re: Addition of Salt

Variation Source	Degree of Freedom	Mean Square	F
Salt Information	1	23.5102	25.36**
Tasters	48	1.2483	1.35
Error	48	0.9269	----
Total	97	----	

**Significant at 1%

Table 12

Effect of Information re: Addition of Sugar

Variation Source	Degree of Freedom	Mean Square	F
Sugar Information	1	9.3462	10.71**
Tasters	38	1.4784	1.69
Error	38	0.8725	----
Total	77	----	

**Significant at 1%

Table 13

Effect of Information re: Addition of Lemon Juice

Variation Source	Degree of Freedom	Mean Square	F
Lemon Juice Information	1	5.6539	5.26*
Tasters	38	1.1545	1.07
Error	38	1.0749	----
Total	77	----	

*Significant at 5%

Table 14

Effect of Information
Mean Scores of Duplicate Samples to the Reference

Type of Information	Mean Score	Number of Tasters
High quality raw product	4.37	132
Low quality raw product	3.71**	132
Chemical treatment	4.02	132
No chemical treatment	4.02	132
Exposed to rays	4.16	132
Not exposed to rays	3.90	132
MSG added	4.27	49
MSG not added	4.14	49
Salt added	4.69	49
Salt not added	3.71**	49
Sugar added	4.41	39
Sugar not added	3.72**	39
Lemon juice added	4.36	39
Lemon juice not added	3.82*	39

**Significantly different at 1%
*Significantly different at 5%

Summary

The results of these flavor tests indicate that some information affects the panel preference while other information does not. The preference judgment of the taste panel was altered by information concerning the quality of the raw product, the presence or absence of salt, of sugar, or of lemon juice in the tomato juice. It was unaffected

by information concerning the presence of monosodium glutamate, treatment with unspecified chemicals, or treatment by rays. This could imply the possibility that only information which has meaning within the range of experience of the tasters will influence the preference results.

It might reasonably be considered that the college students composing these panels had little or no experience with chemical treatments, monosodium glutamate, or sterilizing rays in food products, whereas the use of salt, sugar, or lemon juice in tomato juice or similar products is probably relatively common experience to all. It is likely also that tomatoes as a raw product in varying degrees of perfection or imperfection are well known to these tasters.

These possibilities suggest that if the information conveyed to tasters has meaning within terms of experience, it may influence their preference, but if the information does not have meaning it may not affect the preference judgment. It leads to the consideration that tasters may taste what they expect to taste. It is clear from these flavor tests that information conveyed to, or possessed by, tasters has the potential of influencing the preference findings.

THE EFFECT ON FLAVOR PREFERENCE OF RINSING WITH TAP WATER BETWEEN TASTES

Procedure

To determine the effect of rinsing the mouth with tap water between tastes, six pairs of one ounce samples of tomato juice were presented in random order to each taster of a panel of 134 college students. One cup of each of the pairs was identified as reference sample by R marked on the outside of the cup. The other cup was identified only by a four-digit code number. Three of the six pairs of samples consisted of identical tomato juice. The coded cup of the other three pairs contained the same tomato juice as the identical pairs, while the reference cup of these three pairs contained tomato juice of a different brand which preliminary tests had determined to be of a different preference.

A preference ballot containing specific instructions concerning the rinsing of the mouth with tap water was issued with each pair of samples. Three types of ballots were constructed. Ballot one, shown as Exhibit C, instructed the taster to be sure to rinse his mouth with tap water between tasting the reference and the sample. Ballot two, shown as Exhibit D, instructed the taster not to rinse his mouth with water between tasting the reference

Exhibit C

-- Flavor Test Ballot --

Date:_____ Product:_____ Name:_____

INSTRUCTIONS:

The cup labeled "R" contains a reference sample. Compare the flavor of the numbered sample directly with the flavor of the reference sample.

(a) Taste the reference and the sample as many times as necessary, but Be Sure to Rinse Mouth with tap water between each tasting of R and the sample.

(b) Write the number of the sample in the ballot space which indicates your judgment.

Much better flavor than reference _____

Moderately better flavor _____

Slightly better flavor _____

Neither better nor poorer flavor _____

Slightly poorer flavor _____

Moderately poorer flavor _____

Much poorer flavor than reference _____

BE SURE TO RINSE BETWEEN TASTES

Comments:

When finished, please open sliding door, push cups through. Place ballot in ballot box. Thank you.
BE SURE YOUR NAME IS WRITTEN ON THIS BALLOT

Exhibit D

-- Flavor Test Ballot --

Date:_____ Product:_____ Name:_____

INSTRUCTIONS:

The cup labeled "R" contains a reference sample. Compare the flavor of the numbered sample directly with the flavor of the reference sample.

(a) Taste the reference and the sample as many times as necessary, but Do Not Rinse Mouth with water between tasting R and sample.

(b) Write the number of the sample in the ballot space which indicates your judgment.

Much better flavor than reference _____

Moderately better flavor _____

Slightly better flavor _____

Neither better nor poorer flavor _____ DO NOT RINSE

Slightly poorer flavor _____

Moderately poorer flavor _____

Much poorer flavor than reference _____

Comments:

When finished, please open sliding door, push cups through. Place ballot in ballot box. Thank you.
BE SURE YOUR NAME IS WRITTEN ON THIS BALLOT

and the sample. Ballot three, shown as Exhibit E, instructed the taster to use his normal tasting procedure in arriving at his preference judgment.

Normal tasting procedure with respect to rinsing was ascertained by a written poll of the tasters obtained from each immediately following his tasting. The poll asked the question, "Do you rinse with water between tastes in your normal procedure of tasting tomato juice?," to which in most cases the answer was a direct "yes" or "no." In a few cases the tasters responded with "sometimes" or "yes and no." One hundred five tasters of this panel of 134 stated they did not normally rinse in tasting tomato juice; twenty stated they did normally rinse in tasting tomato juice.

These tests were conducted in individual tasting booths lighted uniformly by white incandescent light. Each was equipped with a stool, a service counter, a pencil, paper rinsing cups, and an individual porcelain basin having hot and cold water supply. Tap water for rinsing was turned on in advance of the tests, was left running during the tests, and was maintained at a temperature of fifty to fifty-five degrees Fahrenheit throughout the test by checking at intervals with a thermometer.

Exhibit E

-- Flavor Test Ballot --

Date:_____ Product:_____ Name:_____

INSTRUCTIONS:

The cup labeled "R" contains a reference sample. Compare the flavor of the numbered sample directly with the flavor of the reference sample.

(a) Taste the reference and the sample as many times as necessary, in whatever procedure of tasting is normal or natural to you.

(b) Write the number of the sample in the ballot space which indicates your judgment.

Much better flavor than reference _____

Moderately better flavor _____

Slightly better flavor _____

Neither better nor poorer flavor _____

Slightly poorer flavor _____

Moderately poorer flavor _____

Much poorer flavor than reference _____

YOUR NORMAL TASTING PROCEDURE

Comments:

When finished, please open sliding door, push cups through. Place ballot in ballot box. Thank you.
BE SURE YOUR NAME IS WRITTEN ON THIS BALLOT

Analysis and Results

The scores obtained from the tasters for the six coded samples were analyzed statistically by analysis of variance on the basis of a 2 x 3 x 13^4 factorial design. This enabled determination of whether the panel preference differed as a result of the three rinsing variables in relation to either of the references. It also permitted determination of whether the scores were different with respect to the two references.

The analysis of variance, shown in Table 15, revealed that the mean scores of the coded samples with relation to the two references were significantly different. The mean scores seen in Table 16, were not significantly different, however, with respect to rinsing, not rinsing, or normal rinsing procedure. Some individual tasters differed in their preference judgments, and tasters reacted differently with respect to the two reference samples, as indicated by the significant interaction between tasters and reference. On the other hand, rinsing procedures uniformly lacked effect on tasters as evidenced by non-significant interaction between tasters and rinsing procedures. Likewise, rinsing procedures had no effect irrespective of the reference, as shown by no significant interaction between reference and rinsing.

Table 15

Effect of Rinsing Between Tastes

Variation Source	Degree of Freedom	Mean Square	F
Reference	1	22.3334	9.34**
Rinsing	2	1.2028	1.09
Reference x Rinsing	2	2.4042	2.17
Tasters	133	2.2739	2.05**
Tasters x Reference	133	2.3910*	2.16**
Tasters x Rinsing	266	1.2817	1.16
Error	266	1.1085	----
Total	803	----	

**Significant at 1%. *Error term for Reference.

Table 16

Effect of Rinsing
Mean Scores of Coded Samples--(134 Tasters)

Performance	Sample Identical to Reference	Sample Different from Reference
Rinsing between tests	4.10	4.37
Not rinsing between tests	4.12	4.30
Normal rinsing procedure	3.84	4.38
Means per reference	4.02	4.35

Summary

It is clear from the analysis that rinsing or not rinsing with tap water did not influence the preference findings of this panel for the tomato juice tested. This result was obtained for samples which were identical to

the reference as well as for samples judged by the panel to be different than the reference. In the absence of interaction between the reference samples and the rinsing procedures, it is evident that rinsing or not rinsing had no effect on panel preference irrespective of the two types of reference used.

THE EFFECT ON FLAVOR PREFERENCE OF SWALLOWING THE PRODUCT

Procedure

In order to determine the effect of swallowing the product on findings of the flavor preference panel, four pairs of one ounce samples of tomato juice, in small paper cups, were presented to each taster of a panel of 123 college students. The tasters were seated for the test in individual taste testing booths which were lighted uniformly by white incandescent lights and equipped with individual porcelain basins. Each pair of samples was served on a metal tray passed to the taster through a sliding-door opening which was then closed during the tasting.

One sample of each pair was identified as a reference sample by "R" marked on the cup. The other sample of each pair was identified only by a four-digit code number on the cup. Two pairs of the four pairs presented contained identical tomato juice. The reference sample of each of the other two pairs consisted of tomato juice of a different brand than the accompanying coded sample. The coded samples in this case consisted of the same juice as that of the identical pairs.

A printed ballot containing specific instructions was

given to each taster with each pair of samples. These ballots were of two types. One of them, shown as Exhibit F, instructed the taster to be sure to swallow the juice each time the reference and the sample were tasted. The other ballot, shown as Exhibit G, instructed the taster to place the juice in his mouth, then to discard it from the mouth into the basin before making his judgment. The ballot in each case instructed the taster to indicate his preference judgment by writing the number of the coded sample in an appropriate space in the column provided.

One of the two types of ballots was presented with one of the identical pairs, while the other type of ballot was presented with the other identical pair. Likewise, the two types of ballot were presented with the two non-identical pairs of tomato juice samples. The order in which the four pairs were presented to each taster was randomized.

Exhibit F

-- Flavor Test Ballot --

Date: _____ Product: _____ Name: _____

INSTRUCTIONS:

The cup labeled "R" contains a reference sample. Compare the flavor of the numbered sample directly with the flavor of the reference sample.

(a) Taste the reference and the sample as many times as necessary, but BE SURE TO SWALLOW this juice each time you taste the reference and the sample.

(b) Write the number of the sample in the ballot space which indicates your judgment:

Much better flavor than reference _____

Moderately better flavor _____

Slightly better flavor _____

Neither better nor poorer flavor _____ SWALLOW THE JUICE

Slightly poorer flavor _____

Moderately poorer flavor _____

Much poorer flavor than reference _____

Comments:

When finished, please open sliding door, push cups through. Place ballot in ballot box. Thank you.
BE SURE YOUR NAME IS WRITTEN ON THIS BALLOT

Exhibit G

-- Flavor Test Ballot --

Date:_____ Product:_____ Name:_____

INSTRUCTIONS:

The cup labeled "R" contains a reference sample. Compare the flavor of the numbered sample directly with the flavor of the reference sample.

(a) Do Not Swallow this juice when you taste the reference and the sample. Each time you taste this reference and this sample, place a portion of the juice in the mouth, Do Not Swallow It, then expectorate it from mouth into basin.

(b) Repeat this procedure as many times as desired, then write the number of the sample in the ballot space which indicates your judgment:

Much better flavor than reference _____

Moderately better flavor _____

Slightly better flavor _____

Neither better nor poorer flavor _____ DO NOT SWALLOW

Slightly poorer flavor _____

Moderately poorer flavor _____

Much poorer flavor than reference _____

Comments:

When finished, please open sliding door, push cups through. Place ballot in ballot box. Thank you.
BE SURE YOUR NAME IS WRITTEN ON THIS BALLOT

Analyses and Results

The scores obtained from each taster were tabulated on the basis of swallowing or not swallowing the identical and the non-identical samples, designated as samples K and samples A, respectively. These scores were then statistically analyzed for variance.

The analysis of variance, seen as Table 17, revealed that scores for sample K and sample A were significantly different when swallowing, but not significantly different when not swallowing. This indicates that when the product was swallowed the tasters were discriminating between the samples, but when it was not swallowed the tasters did not discriminate. The mean scores are shown in Table 18.

Table 17

Effect of Swallowing the Product

Variation Source	Degree of Freedom	Mean Square	F
Tasters	122	1.9811	1.55**
Swallow vs. Not Swallow	1	1.8293	1.43
Swallowing, K vs. A	1	6.8334	5.33*
Not Swallowing, K vs. A	1	0.9147	0.71
Error	366	1.2820	----
Total	491	----	

**Significant at 1%. *Significant at 5%.

Table 18

Effect of Swallowing
Mean Scores of Coded Samples--(123 Tasters)

Samples	Swallowing	Not Swallowing
K (identical to R)	4.15	4.26
A (different from R)	4.49*	4.14

*Difference significant at 5%

Inasmuch as the true value for the K samples is known to be 4, the t-test was applied to ascertain whether the scores obtained for K were significantly different from the true score. The "t" value in the case of swallowing the sample is 1.48, which is not significant. In the case of not swallowing, however, the "t" value is 2.56, which is significant at the five per cent level, signifying that the panel score was significantly different from 4 when not swallowing.

Summary

The results of these tests indicate clearly that swallowing as compared to not swallowing influences the preference findings of the panel for this product. Swallowing of the tomato juice resulted in a true preference score, while not swallowing produced an untrue preference score for samples which were identical to the

reference. Swallowing also enabled the tasters to discriminate between samples which were different. This suggests that in the absence of swallowing, preferences for samples which are different from the reference may also be untrue.

THE EFFECT ON FLAVOR PREFERENCE OF THE QUANTITY OF SAMPLE CONSUMED

Procedure

The quantity of sample consumed as a factor in flavor preference was examined by conducting two separate tests. In the first of these, three pairs of tomato juice samples in paper cups were presented to each taster of a panel of 135 men and women college students. In the second test, 67 men and women college students composed the panel. The tasters were seated for the tests in individual taste testing booths which were illuminated uniformly by white incandescent light. Each pair of samples was served on a separate tray passed to the taster through an opening in the booth which was closed by a sliding door during the tasting. One cup of each pair of samples was identified as the reference by R marked on the outside. The other cup was identified only with a four-digit code number chosen from a random number table.

Each cup of reference contained one and one-half ounces of tomato juice. The coded cup in two of the three pairs contained one and one-half ounces of juice, while the coded cup of the third pair contained one-half ounce of juice. In the first test, all tomato juice for all of the cups was identical, obtained from one lot,

prepared by blending the contents of several forty-six ounce cans into a single container for service into the sample cups. In the second test, all reference cups contained tomato juice of one brand, while all coded cups contained juice of another brand. Preliminary tests had determined that these two brands were of different flavor preference.

A ballot containing specific instructions was presented with each pair of samples. These ballots were of two types. One of them, shown as Exhibit H, instructed the taster to be sure to consume all of the juice in the coded sample cup before rendering his preference judgment. This ballot was presented with one of the pairs of samples consisting of one and one-half ounces of juice in both the reference and the coded cup. It was also presented with the pair of samples which consisted of one and one-half ounces of juice in the reference cup and one-half ounce of juice in the coded cup.

The other ballot, shown as Exhibit I, instructed the taster to drink whatever amount of the sample he desired before making his preference judgment. This ballot was presented with one of the pairs of samples containing one and one-half ounces of tomato juice in both the reference and the sample cups. The order in which the three pairs of samples of juice were presented to each taster was

Exhibit H

-- Flavor Test Ballot --

Date:_____ Product:_____ Name:_____

INSTRUCTIONS:

The cup labeled "R" contains a reference sample. Compare the flavor of the numbered sample directly with the flavor of the reference sample.

(a) Taste the reference as many times as necessary, but Be Sure To Drink All of the Numbered Sample before making judgment.

(b) Write the number of the sample in the ballot space which indicates your judgment:

Much better flavor than reference _____

Moderately better flavor _____

Slightly better flavor _____

Neither better nor poorer flavor _____

Slightly poorer flavor _____

Moderately poorer flavor _____

Much poorer flavor than reference _____

DRINK ALL OF THE SAMPLE

Comments:

When finished, please open sliding door, push cups through. Place ballot in ballot box. Thank you.
BE SURE YOUR NAME IS WRITTEN ON THIS BALLOT

Exhibit I

-- Flavor Test Ballot --

Date: _____ Product: _____ Name: _____

INSTRUCTIONS:

The cup labeled "R" contains a reference sample. Compare the flavor of the numbered sample directly with the flavor of the reference sample.

(a) Taste the reference and the sample as many times as necessary. Drink whatever amount of the sample you desire to make your judgment.

(b) Write the number of the sample in the ballot space which indicates your judgment:

Much better flavor than reference _____

Moderately better flavor _____

Slightly better flavor _____

Neither better nor poorer flavor _____

Slightly poorer flavor _____

Moderately poorer flavor _____

Much poorer flavor than reference _____

Comments:

When finished, please open sliding door, push cups through. Place ballot in ballot box. Thank you.
BE SURE YOUR NAME IS WRITTEN ON THIS BALLOT

randomized.

Determination was made of the average amount of juice consumed by each taster on the basis of the ballot instructions to drink whatever amount of the sample he desired. This optional amount was found to average one and one-tenth ounces per taster in the case of each of the tests. Unused juice remaining in the coded cups was collected, measured for total volume at the end of the testing day, and the average amount consumed per taster calculated according to the number of tasters in the panel.

Analyses and Results

The judgments rendered by each taster for the three coded samples, each representing different quantities consumed of the same juice, were converted to numerical scores, tabulated, and analyzed for variance.

The analysis of variance for each of the two tests, shown as Tables 19 and 20, revealed that the effect of the quantity of juice consumed was clearly significant. Mean scores obtained by the two tests are shown in Table 21.

In order to determine which of the quantities consumed produced the most accurate result, a t-test for the difference between the true score and the panel scores was made in the case of the samples identical to the reference.

Table 19

Effect of Quantity of Sample Consumed
Test A--Samples Identical to the Reference

Variation Source	Degree of Freedom	Mean Square	F
Quantity	2	6.4395	5.40**
Tasters	134	1.5182	1.27*
Error	268	1.1932	----
Total	404	----	

**Significant at 1%. *Significant at 5%.

Table 20

Effect of Quantity of Sample Consumed
Test B--Samples Different From the Reference

Variation Source	Degree of Freedom	Mean Square	F
Quantity	2	43.0050	25.65**
Tasters	66	2.7525	1.64**
Error	132	1.6767	----
Total	200	----	

**Significant at 1%

Table 21

Effect of Quantity of Sample Consumed
Mean Scores of Coded Samples

Quantity Consumed	Test A Samples Identical to R	Test B Samples Different from R
0.5 ounce	3.80	3.09
1.1 ounces	4.21	4.61
1.5 ounces	4.13	4.28
Number of Tasters	135	67

This could be done inasmuch as the true score for preference between identical samples is 4.00, on the scale used, representing neither better nor poorer flavor. The true preference score for the non-identical samples was not known.

The "t" values obtained were 2.13 in the case of one-half ounce of juice consumed, significant at five per cent, 2.23 for the average of one and one-tenth ounces of juice consumed, significant at five per cent, and 1.38 for the case of one and one-half ounces of juice consumed, which is not significant. These values indicate that the panel scores for the samples identical to the reference differed significantly from the true score in the case of each of the two smaller quantities consumed. Only in the case of the largest quantity consumed was the score not significantly different from the true value.

Summary

The analyses of these flavor tests indicate clearly that the preference judgment of the panels was influenced by the quantity of tomato juice consumed, whether or not the sample tasted was the same or different than the reference. The correctness of the judgments rendered by the panel for the samples which were different from the reference could not be ascertained in view of the lack of

a known true value. But in the case of the samples which were the same as reference, it is seen that only the largest amount consumed--1.5 ounces by each taster--produced a true panel preference result. This would suggest that in the conduct of such tests care should be taken to see that tasters consume an adequate quantity of sample.

SUMMARY AND CONCLUSIONS

A series of flavor preference tests were conducted with several different panels of college students to ascertain the effects on flavor panel preference of the following psychological and environmental factors: (1) The number of samples presented simultaneously to the taster; (2) The location of the test; (3) Restaurant-like sound accompanying the test; (4) Information possessed by the taster; (5) Rinsing of the mouth with tap water between tastes; (6) Swallowing the product; (7) Quantity of sample consumed by the taster. A reference-type preference ballot was utilized which provided a seven-point, two-directional word-phrase scale. Commercially packed tomato juice served as the tasting medium.

It is concluded that, with respect to preferences for tomato juice by Oregon State College students:

1. More accurate preference results are obtained when only two samples are presented with a reference sample, rather than three samples or four samples presented with a reference.

2. The location of the test with respect to individual taste testing booths, an open room, or an auditorium does not alter the preference findings of the panel.

3. Restaurant-like, generalized sound, in contrast

to quiet conditions during tasting, does not affect the panel preference results.

4. Information possessed by the tasters concerning the product may alter the preferences obtained by the panel.

5. Rinsing the mouth with tap water between tastes has no effect on the preference expressed by the flavor panel for this product.

6. Swallowing the product produces more accurate preference results than not swallowing it.

7. Flavor panel preference is altered by the quantity of sample consumed by the taster. A larger quantity consumed apparently produces truer results than a smaller quantity consumed.

BIBLIOGRAPHY

1. Abbott, O. D., et al. A survey of food preferences of Florida men. Gainesville, Florida, Florida agricultural experiment station, 1952. 26p. (Experiment station bulletin 500)

2. Baten, W. D. Reaction of age groups to organoleptic tests. Food technology 4:277-279. 1950.

3. Beidler, L. M. Our taste receptors. The scientific monthly 75(6):343-349. 1952.

4. Bengtsson, K. and E. Helm. Principles of taste testing. Wallerstein laboratory communications 9:172,176. 1946.

5. Blake, R. R. and G. V. Ramsey. Perception, an approach to personality. New York, Ronald Press, 1951. 442p.

6. Bliss, C. I., E. O. Anderson and R. E. Marland. A technique for testing consumer preferences, with special reference to the constituents of ice cream. Storrs, Conn., Agricultural experiment station, 1943. 20p. (Experiment station bulletin 251)

7. Boggs, M. M. and H. L. Hanson. Analysis of foods by sensory difference tests. Advances in food research 2:219-258. 1949.

8. Clendenning, T. Flavor in confections. II. Methods of evaluation. Manufacturing confectioner 20(2):23-25. 1940.

9. Crocker, E. C. Flavor. 1st ed. New York, McGraw-Hill, 1945. 172p.

10. Crocker, E. C. Flavors and spices. Encyclopedia of chemical technology 6:581-594. 1951.

11. Crocker, E. C. The nature of odor. Tappi 35(9): 169A-175A. 1952.

12. Crocker, E. C. The technology of flavors and odors. The confectioner 34:7-10. January 1950.

13. Crocker, E. C., L. B. Sjostrom and G. B. Tallman. Measurement of food acceptance. Industrial and engineering chemistry 40:2254-2257. 1948.

14. Dawson, E. H. and B. L. Harris. Sensory methods for measuring differences in food quality. Washington, U.S. Government printing office, August 1951. 134p. (U.S. Dep't. of Agriculture. Agriculture information bulletin 34)

15. Dixon, W. J. and F. J. Massey, Jr. Introduction to statistical analysis. 1st ed. New York, McGraw-Hill, 1951. 370p.

16. Dove, W. F. The relative nature of human preference, with an example in the palatability of different varieties of sweet corn. Journal of comparative psychology 35:219. 1943.

17. Ehrenberg, A. S. C. and J. M. Shewan. The objective approach to sensory tests of food. Journal of the science of food and agriculture 10:482-490. 1953.

18. Engel, R. Experimentelle Untersuchungen uber die Abhangigkeit der Lust and Unlust von der Reizstarke beim Geschmackssinn. Archiv f. d. gesamte Psychologie 64:1-36. 1928. (Abstracted in psychological abstracts 3:1051. 1929)

19. Eppright, E. S. Food habits and preferences; a study of Iowa people of two age groups. Ames, Iowa, Iowa agricultural experiment station, 1950. (Iowa agricultural experiment station. Research bulletin 376)

20. Foster, D., C. Pratt and N. Schwartz. Variation in flavor judgments in a group situation. Food research 20(6):539-544. 1955.

21. Gaebe, O. F. A comparative odor and flavor study of eggs stored in avenized and unavenized fillers and flats. U.S. egg and poultry magazine 46: 346-349. 1940.

22. Guilford, J. P. Psychometric methods. 2d ed. New York, McGraw-Hill, 1954. 597p.

23. Hanson, H. L., et al. Sensory test methods. II. Effect of previous tests on consumer response to foods. Food technology 9(2):56-59. 1955.

24. Harper, R. Fundamental problems in the subjective appraisal of foodstuffs. Applied statistics 4(3):145-161. 1955.

25. Helm, E. and B. Trolle. Selection of a taste panel. Wallerstein laboratories communications 9(28): 181-194. 1946.

26. Henning, J. C. Flavor evaluation procedures. Geneva, New York State agricultural experiment station, 1948. 20p. (Technical bulletin 284)

27. Hicks, E. W. Tasting tests. Food preservation quarterly 8(1):1-5. 1948.

28. Hillary, V. J. Flavor research--food firms dig deep to learn what gives products their taste. The Wall Street Journal, Pacific Coast edition, page 1. January 25, 1956.

29. Hopkins, J. W. Precision of assessment of palatability of foodstuffs by laboratory panels. Canadian journal of research, Section F. Technology 24:203-214. 1946.

30. Ishler, N. H., E. A. Laue and A. J. Janisch. Reliability of taste testing and consumer testing methods. II. Code bias in consumer testing. Food technology 8(9):389-391. 1954.

31. Janowitz, H. D. and M. I. Grossman. Gusto-olfactory thresholds in relation to appetite and hunger sensations. Journal of applied psychology 2(4): 217-222. 1949.

32. Josephson, D. V. Flavor origins -- review of chemical mechanisms affecting flavor acceptability of dairy products. Agricultural and food chemistry 2(23):1182-1184. 1954.

33. King, F. B. Obtaining a panel for judging flavor in foods. Food research 2:207. 1937.

34. Kirchner, J. G. The chemistry of fruit and vegetable flavors. Advances in food research 2:259-296. 1949.

35. Kramer, A. Food quality and quality control. In handbook of food and agriculture. New York, Reinhold publishing, 1955. pp.733-750.

36. Laird, D. A. and W. J. Breen. Sex and age alterations in taste preferences. American dietetic association journal 15:549-550. 1939.

37. Langwill, K. E. Taste perception and taste preference of the consumer. Food technology 3:136-139. 1949.

38. Laue, E. A., N. H. Ishler and G. A. Bullman. Reliability of taste testing and consumer testing methods. I. Fatigue in taste testing. Food technology 8(9):387-388. 1954.

39. Lorant, J. and G. Lorant. Taste test procedure. Corvallis, Oregon, agricultural experiment station, Food technology department, 1948. 3 numbered leaves. (Mimeograph)

40. Mackey, A. O. and P. Jones. Selection of members of a food tasting panel: discernment of primary tastes in water solution compared with judging ability for foods. Food technology 8(11): 527-530. 1954.

41. McCammon, R. B., M. S. Pittman and L. A. Wilhelm. The odor and flavor of eggs. Poultry science 13:95-101. 1934.

42. Mental impact of colors in food studied. Food field reporter 23(19):57. September 19, 1955.

43. Meyer, D. K. The stability of human gustatory sensitivity during changes in time of food deprivation. Journal of comparative and physiological psychology 45:373-376. 1952.

44. Miller, P. G., J. H. Nair and A. J. Harriman. A household and a laboratory type of panel for testing consumer preference. Food technology 9(9):445-449. 1955.

45. Moncrieff, R. W. The choice of flavour. Food manufacture 22:113-118. 1947.

46. Moncrieff, R. W. Relative sweetness. I. Flavours, fruit juices and spices review 11(5):5-6,8. 1949.

47. Morgan, C. T. Physiological psychology. 1st ed. New York, McGraw-Hill, 1943. 623p.

48. Neubert, A. M. and G. H. Carter. A single-sample method for foreign flavor detection. Food technology 9(11):572-575. 1955.

49. Peryam, D. R. New manual sets forth standards for sensory testing. Activities report, research and development associates. Quartermaster food and container institute 5(4):263-267. 1953.

50. Peryam, D. R. and N. F. Girardot. Advanced taste-test method. Food engineering 24:58. 1952.

51. Peryam, D. R. and V. Swartz. Measurements of sensory differences. Food technology 4:394. 1950.

52. Pfaffman, C. Gustatory afferent impulses. Journal of cellular and comparative physiology 17:243-248. 1941.

53. Platt, W. Rational methods of scoring food products. Food industries 3:108-111. 1931.

54. Polemis, B. W. Analysis of food preference surveys--an introduction. Activities report, research and development associates, Quartermaster food and container institute 5(2):149-159. 1953.

55. Sjostrom, L. B. Measuring cereal flavors by the flavor profile method. Transactions of the American association of cereal chemists 9(2): 96-103. 1951.

56. Trout, G. M. and P. F. Sharp. The reliability of flavor judgments, with special reference to the oxidized flavor of milk. Ithaca, New York (State). Agricultural experiment station, 1937. 60p. (Its Memoir No. 204)

57. Wagner, K. G. Sprachelichc Grundlagen der Sinnenprufung von Lebensmitteln. VI. Mitteilung: Geruch and Geschmack der Lebensmittel. Zeitschrift fur Lebensmittel-Untersuchung und Forschung 90:36-46. 1950.

58. Webster's New International Dictionary. 2d ed. unabridged. Springfield, Massachusetts, G. and C. Merriam Company, 1948. 3210p.

59. Weckel, K. G. Put your taste buds to work on your flavor problems. International association of milk dealers bulletin 34:136-140. 1941.

60. Woodworth, R. S. Experimental psychology. New York, Henry Holt and company, 1938. 889p.