Examination of one aspect of consumer perception of four percent butterfat milk

Stanley Bruce Polk

The University of Montana

Follow this and additional works at: https://scholarworks.umt.edu/etd

Let us know how access to this document benefits you.

Recommended Citation
Polk, Stanley Bruce, "Examination of one aspect of consumer perception of four percent butterfat milk" (1973). Graduate Student Theses, Dissertations, & Professional Papers. 2280.
https://scholarworks.umt.edu/etd/2280
AN EXAMINATION OF ONE ASPECT OF CONSUMER PERCEPTION
OF FOUR PERCENT BUTTERFAT MILK

By

Stanley B. Polk

B.A., Mississippi State University, 1970

Presented in partial fulfillment of the requirements
for the degree of

Master of Business Administration

UNIVERSITY OF MONTANA

1973

Approved by:

[Signatures of approved individuals]

Date May 31, 1973
ACKNOWLEDGMENTS

I am indebted to Thomas J. Steele for "chairing" this paper and for providing the advice and encouragement needed to undertake and to complete this study. To Dr. Bernard J. Bowlen goes a deep appreciation for assistance with this paper and for the many times in the last two years that he has taken a personal interest in both me and my scholastic endeavors. The assistance and materials obtained from the local Great Falls dairies are very much appreciated.

The interest and efforts of Virginia Gilmore aided the accomplishment of this study in many ways. To my wife, Diana, goes a loving appreciation for the countless ways in which she contributed to the completion of this master degree. Finally, I am indebted to the United States Air Force for providing this opportunity as well as the University of Montana.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS .................................. ii
LIST OF TABLES .................................... iv

Chapter

I. INTRODUCTION .............................. 1
   The Need for Consumer Research
   The Need for Consumer Research in Fluid
   Milk Marketing

II. PREVIOUS CONSUMER ORIENTED FLAVOR RESEARCH ...... 10
   A Brief History of Research
   Previous Research and Methodology

III. METHODOLOGY ............................. 33
   The Hypothesis
   Testing the Hypothesis
   Limitations

IV. EXPERIMENTAL RESULTS .................. 48
   Main Findings
   Other Findings
   General Information

V. CONCLUSION ............................... 59
   Summary
   Final Appraisal

REFERENCES ....................................... 62
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Confidence Scale</td>
<td>41</td>
</tr>
<tr>
<td>2. Analysis of Ratings by Brand</td>
<td>49</td>
</tr>
<tr>
<td>3. Analysis of Part Three Confidence Ratings</td>
<td>54</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

The Need for Consumer Research

The economic theory of consumers' choice is based on the assumption that the consumer knows what he buys. He is presumed to be an expert buyer who can appraise the quality of the various goods offered for sale and chooses between them by contrasting, one against the other, the price and quality of each good....

Today, the consumer is no longer an expert shopper.¹

The purchase decision grows more complex each day as the consumer who is more educated, more sophisticated, and more affluent than ever is offered a constantly increasing and changing array of products and services. Fifty years ago most consumer purchase decisions related to basic needs for the majority of Americans. Today, however, an American no longer merely decides to buy salt as he did fifty years ago. He must decide between plain, iodized, garlic flavored, hickory smoke flavored, onion flavored, seasoned, and many other types of salt. Almost every purchase decision is becoming increasingly complex as the

variety of brands and the variety of special types of a pro-
duct continue to grow. Consumers can no longer be familiar
with every product and brand. The typical consumer lacks
the necessary technical skill to evaluate quality for most
major purchases such as dishwashers, automobiles, etc.\textsuperscript{1} Con-
sumers employ decision rules to reduce risk. Brand image,
price, previous experience, store image, peer opinions, com-
parison shopping, government reports, private testing reports,
free samples, guaranties, endorsements, and many other cues
of product quality serve as components of consumer strategies
to avoid losses of time, money, ego, and welfare which result
from the consumer's inability to be familiar with every pro-
duct and service that he may need via personal experience.\textsuperscript{2}
How the consumer makes purchase decisions is of great impor-
tance to marketing. If the marketplace is confusing for the
consumer, the consumer's behavior in the marketplace can be
even more confusing for the marketer.

"As a company tries to find the factors accounting
for strong and weak markets, typical consumer explanations for
both tend to be about the physical attributes of the product."\textsuperscript{3}

\begin{itemize}
\end{itemize}
However, actual physical attributes and physical attributes as perceived by the consumer are often not the same. "Perception is never more than a personal interpretation of information." The personal interpretation of information about the physical attributes of a product is at least in part a function of technical factors, mental readiness, past experience, mood, social, and cultural factors, as well as other variables. Other variables affecting perception vary from authority to authority as well as the relative influential importance of each variable. Engel advances needs and attitudes as two variables of high importance in the perception of product attributes. According to Engel perception may even be selective in that certain stimuli may be ignored while other stimuli are actively sought depending upon the consumer's needs and attitudes. Thus, to the marketer the consumer's perception of the physical attributes of a product may be far more important than the actual physical attributes of a product or service. Indeed, most laymen are not qualified to judge the competence of a physician or the services rendered by him. Nevertheless, some consumers choose a physician whose fees are higher and whose location is more inconvenient than other doctors' because some consumers perceive


2Ibid., p. 292.

cost and location as indices of quality regardless of which physician in reality is superior. More factually, the findings of Allison and Uhl in their study of the influence of beer brand identification on taste perception suggest that beer drinkers cannot discern taste differences among various brands of beer when the beers are not labeled. However, when the beers are labeled as to brand, beer drinkers reported significant taste differences. "Such a finding suggested that the physical product differences had little to do with the various brands' relative success or failure in the market."\(^1\) Brand is often a symbol providing cues to which social class, sex, age group, income group, etc., uses a product.\(^2\) At least this has been proven true for several products as will be discussed in Chapter III.

A paramount problem in the marketing of any product is the determination of why consumers purchase a specific product. Information concerning consumer behavior and consumer motivation enables the marketing manager to adjust product variables to achieve a maximum market for the product or to achieve other marketing goals that the company may have for a product. Effective marketing depends upon information about both the product and the consumer. Most products possess several product variables which may be adjusted by altering the

---

actual physical product attributes or by altering the consumer's perception of the product attributes. In other words, products and consumers can be manipulated in the case of most products to achieve marketing goals if enough marketing intelligence has been amassed about both the product and the consumer. Perceptual psychologists have established long ago that actual characteristics and perceived characteristics are often different. Thus, the perceived attributes of a product may vary from consumer to consumer, while the actual physical attributes remain constant. Actual physical product differences may not be a necessary condition for consumers to perceive a difference. Brand image, price, store image, etc., may be sufficient cues for product differences to be perceived when none are present. "Even if there is no physical difference, if people believe things to be true, then they are true, for them! If consumers like a product, they develop definite attitudes about that product." Perceived product attributes are in some cases much more important in marketing plans than the actual product attributes. Advertising, brand image, price, etc., sometimes play major roles in determining consumer perception of product attributes while actual attributes play

---


only minor roles. Products like Cadillac automobiles and Marlboro cigarettes conjure up strong brand images. Consumers of these products are likely to identify with the image projected by the brand. A marked change in the physical characteristics of a product with a strong brand image may have little effect upon the market since the product's consumer is more influenced by brand image than the physical characteristics of the product. A small change in brand image for such a product may cause a large change in the market. In reality today's marketing manager must view his product as a collection of attributes and determine the relative role that each attribute plays in the consumer's perceptions of this group of attributes as a product. This is to ask what are the salient attributes of a product. Knowledge of the salience of attributes is extremely valuable in formulating any marketing strategy. Determining the extent to which consumers can be influenced by brand image, pricing, advertising, etc., is not easy; but, once marketing intelligence is gathered not only can the firm control the actual product attributes for many products, but the firm can usually exert a large degree of control over the consumer perception of his product. The task of the marketing manager then becomes to formulate strategies which maximize the market by managing product attributes.

To formulate these effective marketing strategies, marketing managers must comprehend the relative importance of both the actual attributes and the perceived attributes of any product. Regrettably the marketing executive cannot go
to a reference shelf and determine the relative importance of actual and perceived attributes for his product. Research in this area is valuable in that it indicates to the firm if a dollar may best be spent on product improvement or on other quality cue-providing information sources. Consumer research benefits not only business, but the consumer as well. By better understanding his behavior in the marketplace the consumer can buy more effectively while the government can better devise laws with which to protect the consumer.¹

The Need for Consumer Research in Fluid Milk Marketing

Except when new products are being introduced or old products are suffering setbacks in the marketplace, the features and attributes of products typically receive scant attention from marketing executives....

Once a product is established in the market, it is likely to be ignored until some danger signal such as declining sales is observed.²

Fluid milk is neither a new product nor an old product that is in danger of failing. Rather, fluid milk is an established product which seems to be taken for granted and neglected by marketing research today just as most well established products are.³ Often an indicator of a product's


³Ibid., pp. 24-29.
accepted status is the difference between retailers' and consumers' perceptions of a product. In today's market, gulfs commonly exist between the consumer's perception and the seller's perception of the importance of product attributes in the purchase decision as well as between the perceptions of the brand images of competitors' products. Fluid milk products are no exception.

In 1971, the average American consumed 259 pounds or almost thirty gallons of fluid milk; yet, in Great Falls, Montana, interviews with an owner and manager, an assistant manager, and a marketing manager each representing different local dairies, revealed that local dairy producers do not agree upon who has what share of the market, whether consumers are brand loyal to fluid milk brands, or what factors are most important in consumer purchase decisions. While two executives felt that people were very brand loyal to milk and that people could very definitely taste differences between milk brands, the third executive believed the consumer not to be loyal to fluid milk brands and incapable of tasting differences between brands. Neither was there much agreement among the executives over which factors are most important in the consumer's decision to purchase milk. For instance, one manager considered shelf display to be by far the most

---


important factor in sales and actual characteristics of the milk brands to be a minor factor. On the other hand, another manager believed quality to be the very most important factor in sales, followed by shelf display. The third manager reversed the order making shelf display the most important like the first manager, but the third manager also considered product quality to be very important like the second manager. Rather than misinformation, the conflicting opinions probably represent an honest difference of opinion resulting from a lack of marketing knowledge and research. Each manager seemed very knowledgeable of basic statistics concerning the sale of his product, such as the ratio of 2 percent butterfat milk sales to 4 percent butterfat milk sales, the ratio of store sales to home delivery sales, etc. The agreement between managers on these points was very close. The disagreements between the opinions of managers points directly to the lack of marketing intelligence and the need for research.

1Robert Hansen, private interview with the owner and manager of Hansen's All Star Dairy, Great Falls, Mont., March, 1973; Robert L. Wolf, private interview with the assistant manager of Meadow Gold Dairy, Great Falls, Mont., March, 1973; Leo Marko, private interview with the assistant marketing manager of Ayrshire Dairy, Great Falls, Mont., March, 1973.
CHAPTER II

PREVIOUS CONSUMER ORIENTED FLAVOR RESEARCH

A Brief History of Research

Early Research

My experience at the National Biscuit Company goes back just a little more than 20 years. My first encounter with flavor was concerned with some work we were doing to improve the flavor of a new variety about to be put on the market. Every sample that was edible was submitted to the laboratory director. He was the "Laboratory Taste Panel." He considered his taste buds to be the keenest in the food field and, regardless of anyone else's opinion, he knew exactly which samples merited further consideration and which should be discarded. He had some very strong dislikes in flavor, one of them being lemon. The slightest trace of lemon was cause for discard. Any sample accepted by the laboratory director was submitted by him to the vice president in charge of laboratories.

This vice president had keen taste buds also and, in addition, he knew exactly what the public would buy. Naturally, he discarded many samples approved by the laboratory director. If perchance a sample received his personal approval, he would submit it to the executive committee. The executive committee was comprised of vice presidents, most of whom were heavy cigar smokers. When they met to discuss a sample they would lay their wet cigars on a convenient ash tray and proceed to decide the acceptability of a prospective new variety.¹

Nabisco's flavor research techniques of the pre-1950's were rather commonplace for the period. Firms in the food and beverage industry relied heavily upon the likes and dislikes of a few executives in key positions who typically qualified as flavor experts because of their positions. Nabisco attributes its success during that time to a conservative policy of avoiding too much flavor and avoiding unpleasant aftertaste. The consumer was seldom consulted in matters of flavor research except by a few firms such as the Kroger Food Foundation which maintained a consumer taste panel. The panel, known as The Homemakers' Reference Committee, began before 1938 and consisted of a minimum of 125 members which Kroger found to be as effective as a 750 member panel and much less costly to maintain. Consumers on the panel made paired comparisons in blind tests of common groceries which were received by parcel post. Pairs consisted of a test product and a proven product, and employed such foods as green beans, catsup, instant coffee, salad dressing, and seasonings. Kroger more recently used the panel to evaluate products for use in private branding. However, even up until the late 1950's such consumer oriented taste panels were not common. Even though taste research advanced quickly after 1950, few consumers were involved.

1Ibid., p. 220.
Consumer involvement in flavor research gained attention in 1948 when two psychologists, N. H. Pronko and J. W. Bowles, Jr., published evidence that consumers could not discriminate among colas on a gustatory basis. In a blind taste test subjects were asked to identify four brands of cola. Three colas were large nationally known beverages (Coca Cola, Pepsi Cola, and Royal Crown Cola) while the fourth cola (Vess Cola) was relatively unknown. The experiment was designed to determine what naming behavior subjects would exhibit when attempting to identify the little known cola. In a second condition all four samples of colas were the same cola, yet subjects exhibited essentially the same naming behavior in both conditions. The experimenters concluded that subjects, "...applied a readily available repertoire of cola-naming responses," and that statistics support the view that subjects could not make taste discriminations on the basis of the actual chemical and physical properties of the colas. A second study soon followed, again a product of psychology, and not a product of marketing research. Pronko and Bowles eliminated Vess Cola, the relatively unknown cola, from the sample and essentially replicated the previous experiment using only the three nationally known colas. The prediction that "...the identifications

---


2 Ibid., pp. 304-312.
would be distributed in an order approximating chance," was verified. The experimenters concluded that subjects "... might do just as well by drawing the names of those beverages out of a hat." The observed naming behavior was attributed to advertising and other forms of culturalization.

If the observed naming behavior was indeed the result of advertising and other forms of culturalization, and subjects cannot identify cola on the basis of taste, then it follows that the naming behavior would be observed regardless of the brands of cola used. Thus, in a third study, three relatively unknown colas were utilized (Hyde Park Cola, Kroger Cola, and Spur Cola). Out of 288 identifications, no identifications were correct. Needless to say the hypothesis was accepted. In a fourth study, Pronko and another colleague, D. T. Herman, repeated the procedure for the fourth time. However, one change was made. Subjects in both conditions were informed that colas were either Coca Cola, Pepsi Cola, or Royal Crown Cola under the theory that subjects might do better if the identities of the colas used in the study were known. For all three colas in the second condition and for Pepsi Cola and Royal Crown Cola in the first

---


2 Ibid., pp. 559-564.

condition, identifications did not vary significantly from chance. Surprisingly, however, Coca Cola was identified with a frequency that yields a statistically significant difference from chance expectancy. The experimenters concluded; "Narrowing his choice apparently permits him to make more strikes, although even in this situation he misidentifies Coca Cola almost as often as he identifies it."¹ The cola studies continued to draw attention as late as 1953 when the study was essentially replicated for the first condition in Lebanon using the three leading colas (Coca Cola, Pepsi Cola, and a local imitation called Champagne Cola). Only Champagne Cola was identified significantly more often than chance would predict at the .001 significance level, even though Champagne Cola is a local imitation of the American colas. However, Pepsi Cola had been introduced into the area only six months earlier and was named the most often which was probably the result of heavy advertising.²

Nine years later in 1962, another cola study appeared with several new twists to the earlier 'Pronko' type study. In this instance six paired comparisons were utilized after the experimenter had conducted a survey to determine if the subjects were familiar with the colas to be used in the study.

Again Coca Cola, Pepsi Cola, and Royal Crown Cola constituted the sample, and the subjects were informed of the brands to be tasted. Each brand was tasted four times. No cola was compared against itself, and each subject was made aware of this fact. The experiment found the frequency of identification for Coca Cola and for Pepsi Cola to be significant at the .01 significance level, while Royal Crown Cola was not identified more often than chance would predict. Moreover, 58 percent of the subjects reported not having tasted Royal Crown Cola in the last six months. "No relationship was found between ability to identify cola beverages and consumption (i.e., number of colas consumed in an average week)."\(^1\)

The findings from this series of six studies may seem on the surface to be somewhat contradictory and to a certain extent they are. However, in the fourth study where the frequency of identification of Coca Cola was significantly higher than chance, still Coca Cola was incorrectly identified almost as often as it was correctly identified. In the sixth study only three out of four correct identifications were required to be included in the data as a success. Still the number of misidentifications far exceeded the number of correct identifications.\(^2\)

The cola research came with the birth of a new era in consumer flavor research. Ten years after the publication


\(^2\)Ibid., p. 359.
of the first cola study the use of large consumer panels for
difference tasting were reported as a recent development in
the brewing industry. One brewery used plant visitors for a
taste panel, while another had founded a 2500 member panel.\(^1\)
Pillsbury, also, had begun consulting the consumers' taste
buds by 1958. In developing a refrigerated caramel nut roll,
Pillsbury surveyed consumers to discover consumer wants, and
then returned to the consumer to taste test the new product.\(^2\)
McCormick and Company introduced a consumer taste test panel
into its taste testing program. McCormick maintained individ­
ual taste experts; panels of experts; panels trained, but not
expert, in taste difference detection; and preference panels.
The large consumer panel provided a standard for calibration
of the other panels. Consumer panels were usually maintained
by research organizations, consulting firms, or advertising
firms. Panels employed blind taste tests and ranged in size
from 25 to 1000 members with a few even larger.\(^3\) Even Esso
began using an expert panel when it discovered that the odor

\(^1\) John B. Bockelmann, "Taste Testing from Viewpoint
of the Modern Brewer," in Flavor Research and Food Acceptance,
sponsored by Arthur D. Little, Inc. (New York: Reinhold Pub­

\(^2\) Guy-Robert Detlefsen, "Development of a Product,
" in Flavor Research and Food Acceptance, sponsored by Arthur
D. Little, Inc. (New York: Reinhold Publishing Corporation,

\(^3\) Richard L. Hall, "Flavor Study Approaches at Mc­
Cormick & Company, Inc.," in Flavor Research and Food Accept­
ance, sponsored by Arthur D. Little, Inc. (New York: Rein­
of waxes and oils directly influenced consumer acceptance.\(^1\)

In addition to the popularizing of consumer consultation in matters of flavor, some quality research was undertaken by industry. Most previous flavor research had been performed by psychologists at universities primarily interested in psychological aspects of human behavior as opposed to the business and marketing aspects of behavior. One notable piece of research on the role of color in identifying sherbet flavors was conducted by McCormick and Company. Researchers were primarily interested in examining the effects of color on flavor perception. Six flavors of sherbet (lemon, lime, orange, grape, pineapple, and almond) were each prepared in the commonly associated color, an inappropriate color, and white or colorless. No other variables were introduced. An experienced large scale consumer panel was convened. Subjects correctly identified the flavors with a high degree of success when the associated color was present. In the white or colorless samples, the subjects' ability to correctly identify flavors deteriorated significantly. "When the sherbet was deceptively colored, only a few judges were able to name the flavor correctly, and the great majority named a flavor usually associated with the color in question."\(^2\) In addition, color was found to greatly influence the subjects' estimation of


\(^2\)Hall, "Flavor Study at McCormick," p. 231.
the flavor strength and quality, despite the use of an experienced non-expert panel. The study concluded that color possessed far greater powers of influence over the perception of sherbet flavor than the actual flavoring of the sherbet.¹ Thus, by the late 1950's consumer oriented flavor research had advanced from the cigar smoke-filled executive meeting room of the pre World War II era to the consumer panel and the use of quality research methodology. More importantly the realization began to occur in marketing that consumer perception of attributes and not the actual product attributes often determines the course of consumer behavior.

After 1960

Early in the 1960's Allison and Uhl reported what is probably the most classic study in the field of consumer oriented flavor research. The original consumer beer-tasting study examined the influence of beer brand identification on consumer taste perception of various brands of beer. In a sophisticated design consumers rated six brands of beer, each brand under two neutral labels in the blind condition and under the real label in the brand known condition. Consumers drank the beer at home under normal conditions and rated each beer on nine qualities on a three point scale while overall quality was rated on a ten point scale. Findings indicated that beer drinkers, in general, could not distinguish among

¹Hall, "Flavor Study at McCormick," pp. 229-233.
various brands of beer in a blind test, or identify 'their' brand in a blind test. However, in the brand known condition, beer drinkers consistently rated their brand higher than other beer brands. Also, the overall ratings for all beers increased considerably in the second condition as well as the ratings for most beer qualities, except bitterness and sweetness. The conclusions of Allison, Director of Marketing Research for National Distillers' Products Company, and Uhl, an assistant professor at the State University of Iowa, included actual marketing ramifications resulting from this piece of cooperative research by business and education.¹

Participants, in general did not appear to be able to discern the taste differences among the various beer brands, but apparently labels, and their associations did influence their evaluations. In other words, product distinctions or differences, in the minds of the participants, arose primarily through their receptiveness to the various firms' marketing efforts rather than through perceived physical product differences. Such a finding suggested that the physical product differences had little to do with the various brands' relative success or failure in the market (assuming the various physical products had been relatively constant). Furthermore, this elimination of the product variable focused attention on the various firms' marketing efforts, and, more specifically on the resulting brand images.²

Several years later another important consumer-oriented flavor study was conducted by McConnell to determine the effects of price on brand loyalty development. The effects


of price on consumer perception of beer attributes were also examined. Subjects were given only price information about three samples of beer which were in reality all the same brand; however, labels had been removed and replaced with new ones which represented brands by the use of neutral letters (M, L, or P). Subjects were told that each test brand was different. Money was taped to the bottles in addition to the price information. The money taped to a bottle represented the difference in cost per bottle between the brand and the most expensive brand, thus simulating the savings that would be realized by choosing a brand other than the most expensive.

The study consisted of twenty-four test periods in which a subject selected a bottle of beer for consumption. Any money attached to the bottle was kept by the subject. The development of brand loyalty was observed as subjects quickly developed preferences despite the lack of any difference between the composition of brands. Only 15 percent reported no difference between the brands. These, of course, took the brand with the most money attached. The other 85 percent developed preferences on the basis of perceived differences resulting from price. One subject remarked "M is a good strong malty beer, but I like L because it is light. Mmm!! P would poison me--make me ill. I couldn't finish the bottle." Brand P, the least expensive, was also reported to be undrinkable. Another subject who had developed a preference for P proclaimed N, the most expensive brand, to be, "...the worst I've had, you could not give it away." Brand L, the medium priced brand, caused
a subject to remark to the experimenter, "You must have an apartment full of L to get rid of—you can't blame people for not taking it." After being told that all three brands were the same beer, one subject loyal to M, the most expensive, withheld the information from his wife who had protested throughout the experiment that the three brands tasted the same. Clearly, at least in this example for 85 percent of the subjects, actual physical product attributes and perceived product attributes of beer are different.

In one of the earliest studies of the development of brand loyalty, Tucker had observed that consumers will become brand loyal even when there is no discernible difference between brands other than the brand itself. Four test brands of bread were labeled L, M, P, and H. Again the test brands were in reality all the same brand. Price did not vary in this study, yet brand loyalty developed. To test the strength of brand loyalty money was attached to loaves other than the preferred brand after the housewife had selected the brand three times in succession. If the subject continued to select her preferred brand, the amount was increased from two cents to three cents and so on until a maximum of seven cents was reached. However, the increasing reward for selecting the non-preferred bread may have been perceived as an indication of inferior quality. "No wonder you put the special on

2 Ibid., pp. 13-19.
brand 'P'. It's the worst of all, remarked one participant, again, illustrating that perceived product attributes may vary radically from actual product attributes depending upon brand image, price, and other factors.

Criticizing the existing body of consumer oriented flavor research for not using a wide diversity of food and beverage products, James K. Maken published a study on the effects of brand preferences upon consumers' taste perception of turkey meat. Turkey not only differs from cola, beer, and bread in basic food type, but turkey is purchased infrequently and usually is consumed by persons unaware of the brand. Similar samples of turkey meat were served to subjects who were instructed that one sample came from a well known nationally advertised brand while the other sample came from a little known brand not distributed or advertised in the test area. In reality, however, both samples were from the same bird; yet, subjects preferred the nationally known brand to the unknown brand by a ratio of more than three to two. In part two of the experiment, tough and tender samples of turkey meat (as indicated on a shear press) were served to the same subjects. Eighty percent preferred the tender meat and of these 63 percent believed the tender meat to be cut from the well known brand and the tough from the little known brand.

Thus, once again perceived product attributes such as brand image may cause the consumer to perceive a taste difference that does not exist.

Two non-flavor studies do bear mentioning here. One, an odor study, is closely related to a flavor study since odor is a component of flavor. Housewives in this study were asked to compare several pairs of silk hose. The hose were identical except for odor. Each of the four pairs carried a different scent including a normally scented pair. Only six of 250 subjects noticed any scent, yet the natural scented pair was preferred by only 8 percent of the sample, while 50 percent preferred a narcissus scent, 24 percent preferred a fruity scent, and 18 percent preferred a sachet scent. Even though the odors were not consciously detected, the scents very significantly influenced the judgments of the quality of silk hose.\(^1\)

Another non-food product for which blind tests similar to taste studies have been performed is adding machines. Two brands of adding machines were rated in three conditions. Under the first condition the two machines were rated in a blind test. The little known machine was rated somewhat superior to the better known machine. In the correct labeling condition, the known brand machine was rated superior. When

the labels were reversed the little known, incorrectly labeled machine was judged to be far superior. Thus, "...even when there is a difference between products, quality ratings can be reversed by product information."¹

The next logical step in researching the relationship between perceived product attributes and actual product attributes involves the examination of multiple quality cue situations, as opposed to the single quality cue indicator research presented thus far. A multiple cue study was performed by Jacoby, Olson, and Haddock. "The general purpose of the study was to examine the unique and interaction effects of three cues; price, composition, and brand image--on perception of brand quality." Subjects were divided into eighteen cells of seventeen subjects per cell. Each cell contained a unique set of variables such as price--present or absent, brand name--present or absent, and composition differences--present or absent for each of three brands of beer. A fourth brand was used for the composition differences--absent condition. Each beer brand had been selected in a pretest of twenty brands of beer for subjects' ability to estimate the price of the beer. An ultra premium brand, a popular priced brand, and an inexpensive brand of beer were used. Composition differences did exert a difference in quality perception, but only for the inexpensive brand. The fake inexpensive brand in the composition absent condition was scored higher in all cells than the actual brand. Brand

image also exerted a significant effect upon quality perception, but again for only one brand, the most expensive, and in a positive direction. Surprisingly, price exerted no significant effect upon quality perception except when price was the only cue available. Another surprising and contradictory finding was that, "Given substantial differences in product composition, and on the basis of taste and aroma cues alone, beer drinkers are able to discriminate quality differences in the expected direction across different brands of beer."¹ The experimenters concluded that price played only a limited role in influencing perception of beer quality when other sources of information such as brand names were present, and that beer drinkers possessed at least a limited ability to discriminate between brands of beer when compositional differences between the brands are large.²

Valenzi and Andrews devised a multiple cue study utilizing price information as one cue and product composition as another. Eversweet butter, Imperial margarine, and Nutley margarine (respectively 95 cents, 45 cents, and 20 cents per pound) were found to be rated significantly different in a blind paired comparison design. Subjects believed that they were rating eighteen different margarines and butters on a nine point scale on the basis of overall quality. However,

²Ibid., pp. 570-579.
each of the three forementioned brands was presented six times with different prices, such that each brand was presented twice to each subject with a high, a medium, and a low price tag. The hypothesis predicted that price and actual product difference would be positively related to quality ratings, but price would account for the larger effect. However, results showed that price accounted for only 4 percent of the quality rating variance, while product difference accounted for 13 percent agreeing very well with the Jacoby, Olson, and Haddock study. Valenzi and Andrews doubted the findings and suggested that the effect of price may have been underestimated due to the use of college coeds who were not frequent purchasers of the products in question and, also due to the low absolute price difference between products.¹ Perhaps the experimenters were correct, for not all cues significantly decrease the value of price information to the consumer judging quality. Enis and Stafford had earlier discovered that when price and store image cues were included in an experimental design, price still strongly affected quality perception.²

Previous Research and Methodology

Methodological Findings

Methodology in consumer-oriented flavor research is critical. Extraneous variables can be very easily introduced


²Jacoby, Olson, and Haddock, "Determinants of Perceived Quality," p. 570.
which may be of considerable significance or of none. Successive pourings of beer from the same bottle will appear different to an expert taster.\(^1\) Clearly, this would be of little concern for most consumer-oriented research. Nevertheless, such a factor is a potential explanation for some variance. Tobacco companies discovered long ago that brand image could radically alter consumer perception of cigarette brand attributes. Personality variables play a large role in cigarette brand choice.\(^2\) Cigarette marketers must decide whether to de-emphasize brand attributes to broaden appeal to several market segments or to emphasize brand attributes to appeal in depth to one market segment. Thus, in some methodological designs even personality variables and self images must be considered. Group influence on ratings is another methodological consideration since the opinion of several evaluators may influence other evaluators to express an opinion contrary to their perceptions.\(^3\)

A methodological question in any consumer oriented flavor research is what ability do consumers possess to judge flavor. Previous research has credited the average consumer with not only the ability to distinguish between certain flavors, and to rank and to score certain flavors; but, also, he

\(^1\)Bockelmann, "Viewpoint of the Modern Brewer," p. 217.


\(^3\)Hall, "Flavor Study at McCormick," pp. 233-235.
has been credited with the ability to suggest possible changes in formulation of a product that might improve its acceptability, and with the ability to indicate the probable success or failure of a product.\(^1\) Other research has demonstrated the ability of consumers to compare food samples to ideally acceptable samples on the basis of taste and "...to estimate the magnitude and direction of change that they would make on one or more flavor dimensions of each sample to increase its acceptability."\(^2\) The consumer's ability to learn flavor discrimination among such foods as wines in one session has been shown. However, only an ability to discern differences was shown. Attempts to teach subjects to identify five different wines failed in both conditions of information about the correctness of judgments and no information. Even subjects given the correct identification of the wine after each sample, failed to learn any more than subjects in the other two conditions.\(^3\)

Methodology for consumer-oriented flavor research has constituted the aim of some research. Taste testing order has


been found to significantly influence outcomes of taste tests for many products. Orders must be rotated on a random basis to avoid such effects.\(^1\) Taste test panels of forty to eighty members have been found to be as effective as panels of 160 persons for rating fruit drinks.\(^2\) Discriminations between samples of accessory foods (catsup, cake icing, etc.) have been shown to be as good without the carrier and in some cases better than with the carrier.\(^3\) The effects of sex and age also have important implications for methodology. "Study of the interactions of sex, age (above ten years of age), preference, and reason for preference revealed no consistent relationship among these variables,"\(^4\) in a taste study conducted on sugar and brine contents in canned peas. A study on consumer preference for shape and flavor of almonds indicated that, "... neither time of day nor differences among days affected preference."\(^5\)


Methodological Implications

Previous research discussed in this chapter has chiefly been directed at goals other than determining the consumer's ability to distinguish between products on the basis of physical attributes. After Pronko's first cola experiment he formulated his second experiment to discover what ability the consumer could demonstrate in identifying well known colas in a blind test on the basis of physical attributes. He found none. However, in a later test when subjects were informed of the brand names which they were to identify, the frequency of correct identifications became significant. Other than Pronko's examination of one aspect of consumer product perception using only physical characteristics for information, most other research in the flavor area has either used physical attributes in a multiple cue situation or examined the effects of various other informational forms upon perception of actual attributes. Implications concerning consumer discriminatory capability have been for the most part only secondary products of research. For a few products such as beer, a profile of the consumer's discriminatory ability has been constructed piecemeal. Perhaps, if the profile had been first developed, then pursuing research could have utilized resources more efficiently rather

---

2 Pronko and Herman, "Cola Beverages: Postscript," p. 69.
than having to develop a particular piece of the profile for use in the research at hand.

The question of the consumer's ability to "taste the difference between" is relative to the degree of identification of differences and relative to the degree of training. Is "tasting the difference between" merely to say that A is different from B, or must the subject identify A as brand "X" and B as brand "Y"? For most food products, taste experts exist who can discriminate differences and identify brands.¹ Yet, the average consumer may not be so trained as the expert. Certainly this is true for beer, but perhaps the consumer of Chateau Rotheschilde would closely rival professional wine tasters. Conflicting conclusions about consumers' ability to identify colas and to discriminate among beer brands may be the results of nothing more than methodology and definitions. Thumin took issue with Pronko's cola findings and used paired comparisons, a much stronger discriminatory tool than Pronko's taste and identify method,² to demonstrate an ability to identify Pepsi Cola as well as Coca Cola. The seeming contradiction may be nothing more than the use of a stronger method to identify a smaller difference than that with which Pronko was concerned.

One thing seems to be clear. Consumer-oriented perception research would benefit by first establishing a profile

¹Dember, The Psychology of Perception, p. 257.
of consumer perceptions of the product's physical attributes before attempting to examine the effects of other variables upon assumed consumer perceptions which must be simultaneously verified.
CHAPTER III

METHODOLOGY

The Hypothesis

The hypothesis for the experiment was: Subjects cannot demonstrate an ability to discriminate among the four major brands of locally produced 4 percent butterfat, homogenized, pasteurized milk on an overall basis of physical attributes in a blind test. The four major brands of 4 percent butterfat, homogenized, pasteurized milk are defined to be the milks commonly available in the Great Falls, Montana area which are processed and distributed by Meadow Gold Dairy, Ayrshire Dairy, Hansen's All Star Dairy, and Vita Rich Dairy. Four percent butterfat is defined to include milks ranging in butterfat content from 3.3 percent to slightly over 4 percent which are commonly known in the industry as 'homo'. Such milks will hereafter be referred to as test milk or 4 percent milk even though the butterfat content is not exactly 4 percent.

Testing the Hypothesis

Design and Intent

The experimental design consisted of three basic tasks designed to indicate the ability or lack of ability of
subjects to make discriminations among the test milks on the basis of taste. The first part of the experiment was engineered to determine if any one brand of 4 percent milk was commonly preferred by the subjects. If any one brand was generally preferred, then it would be quite clear that subjects did demonstrate a capability to discern among the four major brands of 4 percent milk and shared a common preference for one of them. However, the lack of this ability would not be sufficient to conclude that subjects could not discriminate among the test milks. Indeed, subjects might well be able to discriminate among the test milks and register no common preference if each test milk was equally preferred due to its own unique set of characteristics. Perhaps a factor which would allow subjects to discriminate at least one brand of milk from the others would be a factor which does not influence perceived milk quality. The second portion of the experiment controlled for this eventuality.

In the second part of the experiment, subjects were required to make attempts at identifying brands with which they were familiar. Subjects could use only the physical cues of the test milks with which to identify the brands. Still, even if subjects could not demonstrate an ability to discriminate among the test milks by identifying brands, the possibility

remains that differences between the test milks are very small and could be detected only in a direct comparison type test. Such tasks test only for the ability of the subject to detect similarities. The final phase of the experiment was designed to accomplish exactly that task. If none of the parts of the experiment revealed an ability to discriminate by one of the three methods described, then the hypothesis would be accepted. Should any one of the parts of the experiment expose a talent for discerning among the test milks, the hypothesis would be rejected.

Hopefully, the design for the experiment would indicate to some extent the degree of ability subjects have to discriminate among the test milks should any of the three parts of the experiment yield significant results. For instance, if part one of the design demonstrated that one brand was commonly preferred, then that finding would indicate a consumer ability of greater magnitude than the abilities that could be demonstrated in the similarity/dissimilarity testing of part three. In other words, test three is a much more powerful test than the test of parts one and two. It is obvious that the expert tasters do exist and that their skill is learned. One professor estimates that 90 percent of his dairying students, "...can be trained to become effective judges of dairy products, and around 10 percent of them can be developed into superior judges."¹ Since most Americans

consume large quantities of milk during their lifetimes, the question becomes how much ability to "taste the difference" between milk brands is learned on an informal experience basis. Thus, a design that possesses some power to reflect the degree of ability is desirable.

The Method

Part One

Subjects were first asked to rate four samples of milk after having tasted all samples. Only one subject was tested at a time. Samples were presented in plain four ounce paper cups. To the subject all cups were identical. A small piece of masking tape was attached to the bottom of each cup, and on the tape was penciled an identifying letter in such a manner that it was highly unlikely that the subject would be able to observe the identifying letter at any time during the course of the experiment. If he should see the tape, he would see only a neutral letter, K, L, M, or P.\(^1\) A rating scale from one for extremely poor quality to 100 for extremely high quality was devised. Subjects were instructed to taste all samples in a prescribed order before evaluating any sample. After all samples had been tasted at least once, subjects were allowed to taste any sample in any order as often as they liked before rating the samples. Such a procedure overcomes the tendency to prefer the sample tasted first when only

\(^1\) McConnell, "Brand Loyalty," p. 449.
a few samples are employed. The experimenter read the following instructions to each subject before beginning this phase of the experiment.

"Before you are four samples of milk. You are to taste each sample of milk starting on the left and taste each sample in the prescribed order, you may retaste any sample in any order you please. After having tasted all the samples at least once, please rate each sample on a scale of one to 100. One indicates extremely low quality, and 100 indicates extremely high quality. Make your ratings on the basis of overall quality. You may rate the samples in any order you desire retasting any sample as often as you like. Record your ratings on the piece of paper in front of each sample. Be careful to keep the samples in the same physical order throughout the experiment. Are there any questions?"

After the ratings were completed, subjects were given more instructions. "Please indicate which milk of the samples you just tasted seems to be the richest, the highest in butterfat. You may retaste any sample in any order you please." Once the indication was made, the subject was given final instructions for part one. "Please rinse your mouth thoroughly with the water provided."

Only questions pertaining to the performance of the task at hand were answered during the experiment. Most

---

questions asked for instructions and could be answered either
by rereading to the subject a portion of the instructions or
by stating that the question could not be answered at that
time.

Part Two

After the experimenter had recorded the data, the
tray of samples was removed from the subject to a work area
behind a screen and out of the subject's sight. There, samp­
les were refilled and the position order of the samples was
rearranged according to a preplanned system. This portion
of the research sought to determine if there existed any
characteristics of the samples sufficiently unique to permit
identification of the milk brands in a blind test. The design
tested for both unique qualities that were independent of
brand choice as well as unique qualities that might affect
brand choice. Subjects were again read instructions by the
experimenter.

"The four samples of milk before you are Meadow Gold,
Ayrshire, Hansen's All Star, and Vita Rich; although they are
not necessarily in that order. Each sample is the same 4 per­
cent butterfat homogenized, pasteurized, milk commonly avail­
able in the Great Falls area."

"Do you regularly use any one of these four brands
of 4 percent butterfat milk? By regularly, I mean that you
have tasted the brand of milk at least ten times in the last
six months and have used this brand more than any other
brand during that period." (Experimenter records the reply.)
"Please taste the samples of milk again from left to right just as you did before. After tasting the samples in the prescribed order, you may retaste any sample in any order you please and as often as you like. After you have completed your tasting, please identify to me the sample which you believe to be your regular brand. It is not necessary that you be positive about your identification of the brand, but please try your best. Do you have any questions?"

After the identification had been made the subject was told, "If you believe that you can identify the brand of any other of the samples please do so even if you are not positive about your identification of the brand. Do you have any questions?" Questions in part two were answered in most cases by rereading a portion of the instructions or by stating that the question could not be answered at that time. Phase two of the testing was completed when the subject had identified all the samples which he felt that he might be able to identify correctly. He was then instructed: "Please rinse your mouth thoroughly with the water provided."

Part Three

With the recording of the data from part two, the experimenter again removed the tray, refilled the samples, rearranged the position order of the samples in accordance with the preplanned scheme, and added a fifth identical cup. The new cup was placed in front of the row of the samples in the same place for each subject. The fifth cup of test milk was also nearest the subject and had been filled with one of
three test milks for each subject. Each test milk was used one third of the time except for Vita Rich. The experimenter read the instructions.

"Again before you are samples of Meadow Gold, Hansen's All Star, Ayrshire, and Vita Rich 4 percent butterfat milk. The position order of the samples has been randomly rearranged. A fifth sample has been placed in front of the other four samples. You are to match the fifth sample with the one of the four samples which is the same brand as the fifth sample. Again, please taste the samples left to right, then taste the fifth sample last. Once you have tasted all five samples in the prescribed order, you may retaste any sample as often as you like in any order. Inform me when you have matched the sample. Do you have any questions?"

Again, questions were answered only as they concerned the task at hand and for the most part questions were answered by rereading portions of the instructions. The third and final phase of the experiment examined the ability of subjects to detect differences and similarities among the test milks. If the subject could "taste the difference" then he should be able to match the milk samples on the basis of similarities and dissimilarities.\(^1\) After the subject had matched the sample, the results were recorded by the experimenter.

---

\(^1\)Peryam, "Sensory Difference Test," pp. 50-56.
Follow-Up

The subject then was told that the experiment was concluded and was asked a few questions. The experimenter then recorded the answers to the following questions.

1. On a scale of one to nine as shown to you (experimenter shows the subject Table 1), how much confidence do you place in your identification of the sample you identified in the second part of the experiment as the brand you regularly drink?

<table>
<thead>
<tr>
<th>Rating</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.</td>
<td>Extremely certain</td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Maybe right, maybe wrong</td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Extremely uncertain</td>
</tr>
</tbody>
</table>

2. (If the subject identified more than one sample in part two, the experimenter asked this question.) How much confidence do you place in your identification of the second (third, fourth) sample you identified in part two? Again use the scale shown to you.

3. On the same scale of one to nine, how much confidence do you place in your matching of the fifth sample with the other sample of the same brand in the last part of the experiment?

4. What do you believe the purpose of this experiment to be?
5. What is your age?
6. Are you married?
7. How many children do you have?
8. How much milk does your family buy per week—both 2 percent butterfat and 4 percent butterfat?
9. How many glasses of 2 percent and 4 percent butterfat milk do you drink per week?
10. Do you believe that you can taste the difference between brands of 4 percent butterfat milk?
11. Do you smoke? If so, what do you smoke and how much per day?

"Please do not discuss this experiment for at least two weeks. Knowledge of the experiment could prejudice the performance and judgments of other people who will be participating in the experiment. Your cooperation is appreciated very much. Again, thank you for taking time to participate. Once the experiment is concluded I will be more than happy to discuss the findings with you." The subject was then thanked again but in a much more personal manner. Questions were discouraged when possible or the subject was told that the question could not be answered at that time but that the experimenter would be glad to answer the question once the experiment test periods were completed.

Controls

Volunteers to serve as subjects were solicited from military personnel and dependents as well as a few civilians who either responded to signs posted in the Malmstrom Air Force Base education building or were personally asked by the experimenter to participate while in the education building.
Subjects were tested on a time available basis on Tuesday the 27th of March, 1973; Wednesday the 28th; or Friday the 30th, between the hours of 7:45 a.m. and 5:30 p.m. Room 224 of the education building served as the experimentation room. A partition separated the table at which the subject sat from the work area of the experimenter, allowing the experimenter to prepare samples, to arrange sample orders, and to accomplish any task he desired without the subjects' knowledge of what the experimenter was doing. Samples used were all fresh samples of each brand obtained directly from the dairy and held for no longer than fifty-two hours before being used in the experiment. Milk samples were stored in a refrigerator at 35 degrees F. to 39 degrees F. The temperature of the refrigerator was checked several times each test day. Samples were served directly from the refrigerator in plain white four ounce cups. To insure that samples remained at the same temperature, all five samples were poured at the same time. If the subject drank more of one sample than of the others, all samples were reduced to approximately equal amounts to prevent the subject from identifying one sample from the others from one part of the experiment to the next. Also, this procedure insured that sample temperatures would be very close to the same between all samples when cups were refilled. Milk cartons were opened only several hours at the very most before the final contents were used. Opened cartons were discarded at the end of each test day and fresh samples were obtained on Monday the 26th and Thursday the 29th. Each opened carton
of milk was examined by the experimenter to insure that no gross irregularities such as spoiled or rancid milk were present. Samples were then placed on a tray for transportation to the subject at the table. The twenty-four possible position orders of samples had been drawn in advance from a hat which contained each order twice. This was repeated three times to yield three different lists of forty-eight positions containing each possible position twice. Thus, all possible orders were equally represented but in a random fashion. One list was used for part one of the experiment, the second list was used for part two, and the third list was used in the third part of the experiment. By using three lists the hazards of repeating one list three times were avoided. For the forty-ninth subject the orders were the same as for the first subject, and so on.

Subjects returned the samples to the same position on the tray after each tasting. The experimenter watched each subject closely to insure that the same order was maintained throughout each phase of the design. Masking tape was used to divide the back portion of a standard size cafeteria tray into four adjacent four inch squares to facilitate preservation of the assigned position order throughout each phase of the experiment. To standardize the instructions to subjects throughout the test, the experimenter read the instructions to each subject and attempted to answer all questions concerning the experiment by rereading appropriate portions of the instructions.
Analysis of Data

Data resulting from the ratings of milk samples in part one of the experiment is almost ideally suited for one way analysis of variance. After computing the mean score for each sample, the application of one way analysis of variance indicates "...whether the discrepancies among the means may reasonably be attributed to chance or whether they are indicative of differences among the means of the corresponding populations."¹ In other words, is the variance between the means of ratings for milk samples the result of chance or the result of the milk samples being rated differently? The .05 level of significance was chosen. Calculation of the one way analysis of variance was performed using a pre-stored program in the Honeywell 600 RADC computer.

The much simpler chi squared test statistic functioned to test for significance of the results of both phases two and three of the experiment. In each case the probability of correctly identifying or correctly matching a sample was .25. Thus, the determination of the frequency of expected correct judgments is made by multiplying .25 times the number of subjects. The data yields the frequency of correct judgments observed which constitutes the final information needed for the chi squared sample statistic.² Again a significance level

of .05 was considered acceptable. The number of attempts to identify a second or a third brand of milk in part two of the experiment was so small that no statistic was employed to test this data due to the very small sample size. The pre-stored Honeywell 600 computer program for T testing confidence ratings was used. The T test was used to determine if the differences between confidence ratings of subjects who were correct and who were incorrect were significant.

Limitations

Admittedly the design is limited in that it is artificial. Most consumers do not perform a blind side by side comparison of milk brands in forming attitudes about the brands. Circumstances surrounding the evaluation of milk brands in the home vary greatly from circumstances in this experimental design. Yet, the design is capable of indicating an ability to discern differences among milk brands. Should such an ability not be demonstrated, methods used in the home for evaluating the physical qualities of various brands of milk are of little consequence for the hypothesis in question.

The design is also limited in that it does not employ the professional tasting techniques of tasting the milk sample, spitting, rinsing with water, spitting, and then waiting at least one minute before tasting the next sample. ¹ Such an addition to the design would increase the time required for each subject from fifteen minutes to at least thirty minutes.

¹Wolf, private interview.
or more. Other than the aesthetic drawbacks, the incorporation of such a technique would make the design even more artificial.

A more powerful direct comparison taste test could have been employed in part three by having subjects judge paired comparisons as being the same or as being different. However, such a method requires several paired comparisons which increases the time needed to test one subject. The matching task employed in part three is almost as powerful and is more relevant to marketing in that the matching task requires the subject to identify from a field of four, the one sample which is the same, rather than merely proclaiming that the subject can or cannot taste a difference.

The sample of the consumer population may also be a limiting factor, in that a disproportionate number of males were included. However, value judgments about subject populations can only be speculation unless it is known who makes the decision of which milk brand to purchase for the family and how this decision is made. Other studies suggest that age and sex make no difference in subject preference at least for some foods.¹

CHAPTER IV

EXPERIMENTAL RESULTS

Main Findings

Subjects were not able to demonstrate an ability to discriminate among the four brands of test milks in any one of the three parts of the experiment at the .05 significance level. The hypothesis was accepted.

Ratings for Ayrshire, Meadow Gold, and Vita Rich were almost identical in part one of the experiment. A one way analysis of variance about the means for the ratings of these brands indicated that the probability of the variance between samples' ratings occurring by chance was 91 percent. Ratings for Hansen's All Star averaged 6 to 7 points lower than the other 3 brands. When Hansen's was included in the one way analysis of variance about the means, the probability that the variance among the ratings of the 4 brands was due to chance, dropped to 22.8 percent. The probability that the variance between Hansens' ratings and the ratings of the highest rated brand, Meadow Gold, was due to chance was only 6.3 percent--almost significant at the .05 level. Variance between the ratings of all brands other than Hansen's, was probably the product of chance, (see Table 2).
### TABLE 2

**ANALYSIS OF RATINGS BY BRAND**

<table>
<thead>
<tr>
<th>Brand</th>
<th>Mean</th>
<th>Probability that Variance Between Ratings is Due to Chance</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Ayrshire</td>
<td>Hansen's</td>
</tr>
<tr>
<td>Ayrshire</td>
<td>73.89</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Hansen's All Star</td>
<td>67.14</td>
<td>.114</td>
<td>1.000</td>
</tr>
<tr>
<td>Meadow Gold</td>
<td>74.89</td>
<td>.777</td>
<td>.063</td>
</tr>
<tr>
<td>Vita Rich</td>
<td>73.28</td>
<td>.854</td>
<td>.171</td>
</tr>
</tbody>
</table>
Subjects did not demonstrate an ability to judge butterfat content. Ayrshire (3.75 percent butterfat) was judged to be the highest in butterfat 20 times; Meadow Gold (3.325 percent butterfat) was judged to be the highest in butterfat 19 times; and, Hansen's (3.4 percent butterfat) was judged to be highest in butterfat 17 times. Vita Rich (3.3 percent butterfat) was judged highest in butterfat only 10 times, yielding a chi square of 4.2857. With 3 degrees of freedom, the ranking would be significant at approximately the .25 level. Vita Rich is highly filtered to produce a constant flavor year around. Due to filtering Vita Rich seldom suffers from such off flavors as 'cowy' flavors, 'feedy' flavors, wild onion flavors, etc. According to the Great Falls area manager of Vita Rich, some people object to the lessened effect of some natural milk flavors which are also filtered out.\(^1\) The difference in flavor resulting from filtering rather than a lower butterfat content, may account for the observed less frequent choice of Vita Rich as the highest in butterfat. Especially since Ayrshire, by far the highest in butterfat, was identified as the highest in butterfat only one more time than Meadow Gold which has a butterfat content very close to the butterfat content of Vita Rich.

Part two of the experiment found 33 subjects who had a regular brand of 4 percent milk as defined in the experiment. Subjects who drank more skim milk or 2 percent butterfat milk than 4 percent milk were not considered. All except one subject used Hansen's as their regular brand. This was the result of using military personnel and dependents. Most subjects purchased their milk at the government commissary which carries only one brand of milk. Hansen's All Star Dairy had supplied the commissary with milk under contract for approximately three months preceding the experiment. Only 9 subjects correctly identified their brand in this blind test. The expected frequency of correct identifications was 8. Thus, the chi square value was very low and not significant. The surprising finding of part two was that 8 of the 9 subjects who correctly identified their brand were smokers. Smokers constituted only 41.41 percent of all subjects, but they constituted 54.55 percent of the participants in part two of the experiment. The frequency of correct identifications for smokers in this phase of the experiment, yields a chi square value of 3.63, which is significant at approximately the .30 level for 3 degrees of freedom. The tendencies of smokers to be more likely to be regular users of 4 percent milk was unexpected, but the tendency to be more apt at identifying their brand in a blind test was contrary to expectations. Seven attempts were made to identify a second brand. Four were correct and 3 of these 4 were smokers, while 4 of the 7 had been correct on their first attempt. The sample is too small
for the use of sample statistics, but it is doubtful that these subjects who attempted a second identification possess any superior ability to identify milk brands since in the third part of the experiment only one of the 7 correctly matched the brands and 4 of these were attempting to match their regular brand.

Subjects did not appear to confuse any particular brand with their brand. That is, all brands were named almost the same number of times as the subject's regular brand even though all but one of the subjects used Hansen's as their regular brand. Subjects should have done just as well by chance had they not even tasted the samples! The average confidence rating for part two was 4.229 for subjects who were incorrect and 4.667 for subjects who were correct. Overall, the average confidence rating was 4.325.

In part three, 69 subjects attempted to match either Ayrshire, Hansen's, or Meadow Gold with another sample of the same brand from samples of all 4 brands. Only 19 correctly matched the samples resulting in another low and non-significant chi square value. Subjects matched all three brands with about the expected frequency of correctness. The most interesting finding was that subjects who correctly matched the brands assigned lower confidence ratings to their decisions than did subjects who incorrectly matched the samples. The average confidence rating for a correct subject was 3.895, while the average confidence rating for an incorrect subject was 5.260. Student's T indicates that the probability of the
ratings of the correct subjects being a random sample of the ratings of all subjects was .0335. The most significant difference between confidence ratings of correct and incorrect subjects was for those subjects who attempted to match Ayrshire samples. The probability that the confidence ratings of subjects who correctly matched Ayrshire occurring by chance from the ratings of all subjects was .0384. Other confidence rating variances are not significant at the .05 level; however, several come close, (see Table 3).

Thus, subjects failed to demonstrate an ability to discriminate among the four brands of locally produced 4 percent butterfat in each of three blind tests. Therefore, the hypothesis was accepted.

Other Findings

Each phase of the experiment produced one unexpected tendency or finding. In part one Hansen's All Star was rated an average of 6 to 7 points below the other three brands, all of which were rated on the average within 1.61 points of each other. Hansen's ratings introduced the majority of the variance as reflected in Table 2. Moreover, during the course of the experiment several subjects demonstrated an ability to identify Hansen's as being different from the other samples. The first subject to report that he could distinguish Hansen's from the other samples, pointed to the Hansen's sample in part one of the experiment and said, "That one has a chalky taste." The experimenter noted that the sample was Hansen's. In the second part of the experiment the subject said, "There's that chalky one, again." The sample was again the Hansen's sample.
<table>
<thead>
<tr>
<th>Brand Matched</th>
<th>Number of Subjects</th>
<th>Average Confidence Ratings</th>
<th>Probability that Sample Occurred by Chance from the Sample of:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>By Correctness</td>
<td>Correct and Incorrect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Only that Brand</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>All Brands</td>
</tr>
<tr>
<td>Ayrshire</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct</td>
<td>6</td>
<td>3.500</td>
<td>.0384</td>
</tr>
<tr>
<td>Incorrect</td>
<td>19</td>
<td>4.894</td>
<td>.0557</td>
</tr>
<tr>
<td>Hansen's</td>
<td></td>
<td></td>
<td>.1596</td>
</tr>
<tr>
<td>Correct</td>
<td>6</td>
<td>4.000</td>
<td>.2562</td>
</tr>
<tr>
<td>Incorrect</td>
<td>16</td>
<td>5.875</td>
<td>.1960</td>
</tr>
<tr>
<td>Meadow Gold</td>
<td></td>
<td></td>
<td>.2252</td>
</tr>
<tr>
<td>Correct</td>
<td>7</td>
<td>4.143</td>
<td>.1946</td>
</tr>
<tr>
<td>Incorrect</td>
<td>15</td>
<td>5.066</td>
<td>.3010</td>
</tr>
<tr>
<td>All Brands</td>
<td></td>
<td></td>
<td>.2285</td>
</tr>
<tr>
<td>Correct</td>
<td>19</td>
<td>3.895</td>
<td>.0335</td>
</tr>
<tr>
<td>Incorrect</td>
<td>50</td>
<td>5.260</td>
<td>.1204</td>
</tr>
</tbody>
</table>
After the conclusion of the experiment, the experimenter asked the subject where the chalky sample had been in the third part of the experiment. Again, the subject was correct, as he was for the fourth time, when in a post experiment test the subject again identified the same sample as being chalky from a new random order. He seemed very confident about each identification. He drank more milk per week than any other subject (over 50 glasses per week—all 4 percent butterfat). Despite the fact that he drank over 50 glasses of milk per week, all of which were Hansen's, and despite the fact that he could distinguish "the chalky one" from other samples in a blind test, he did not identify "the chalky one" as Hansen's in part two of the experiment, nor did he at any time show any sign of associating the brand Hansen's with "the chalky one." When he was informed of the identity of the chalky sample, he seemed surprised. Two other subjects demonstrated a similar ability, while two others identified Hansen's as being "most unlike the others," but could not repeat the performance in a post experiment test. To insure that the Hansen's samples were not typical of Hansen's milk due to some irregularity that was present in the milk produced during the week from which the test samples were obtained, the owner manager of Hansen's All Star Dairy was contacted. He knew of no differences present in the milk produced from which the samples were obtained, and he knew of no reason for some subjects being able to detect a difference between his brand and other brands. Since the ability to identify one milk brand as being most
unlike the others was not the object of any test, it is possible that other subjects detected the difference and did not report it. However, subjects did no better at matching Hansen's samples in part three of the experiment than did subjects who were to match other brands. (Unfortunately, all five of the subjects who reported a difference in the Hansen's sample attempted to match samples other than Hansen's in part three of the experiment.) These findings suggest that most subjects did not possess the ability to identify Hansen's as being different. It seems that only a few subjects possessed an ability to detect this difference. This may account for Hansen's overall lower rating since the 5 subjects who demonstrated the ability assigned an average rating of only 37.6 to Hansen's, 50.4 to Vita Rich, 76.6 to Meadow Gold, and 85 to Ayrshire. Assuming that several other subjects detected the difference, but did not report it, and assuming that they rated the milks similar to those subjects who did report the difference, the magnitude of the Hansen's rating would be expected. However, this is only an unverified potential explanation. More importantly, 4 out of 5 of the subjects who detected and reported the difference used Hansen's 4 percent milk regularly; yet, none identified the different sample as Hansen's either in part two of the experiment or at any other time! This fact suggests that the unique quality of Hansen's detected by the subjects was noticeable only in a direct comparison of Hansen's with other brands or that the unique quality was transient.
The tendency of smokers to identify their milk brand correctly with a much greater frequency than non-smokers in part two of the experiment may be the results of the small sample size. On the other hand, some schools of psychology suggest that the smoker is much more sensitive to oral sensation than the non-smoker, which is part of the reason he smokes. Given an increased sensitivity to oral stimuli, smokers should exhibit a higher frequency of correct identifications than non-smokers. Again in part three of the experiment, smokers' 39.29 percent frequency of correctness was much higher than the 19.51 percent rate for non-smokers. However, the difference is not significant at the .05 level. Due to the nature of chi square, if the 39.29 percent rate of correctness for smokers continued as the sample size of smokers increased to 65, then the chi square value would increase to significance at the .05 level. Also, the same is true for smokers in part two. Thus, a larger sample of smokers could verify or dismiss the observed tendency.

The last unexpected finding was significant at the .05 level. In part three, subjects who correctly matched the samples, consistently placed less confidence in their decisions on the average as reflected by confidence ratings, than subjects who were incorrect. No explanation is offered.

General Information

The sample consisted of a total of seventy subjects ranging in age from 16 years to 58 years. Thirty-three
subjects were in their twenties, 20 subjects were in their thirties, 9 subjects were in their forties, 6 were teenagers, and 2 were over 50 years of age. Of these, 9 were females, and 5 were Negroes.

Subjects ranged in their consumption of 4 percent milk from no glasses per week, to over 50 glasses per week. Twenty-three drank 7 or more glasses of 4 percent milk per week. The frequency of smokers drinking 7 or more glasses of milk per week was not significantly different from the overall sample. The frequency of correctness in parts two and three was not significantly different from the overall sample for subjects who drank 7 or more glasses of 4 percent milk per week.
CHAPTER V

CONCLUSION

Summary

Subjects could not demonstrate an ability to discriminate among the four brands of locally produced 4 percent butterfat, homogenized, pasteurized, milk in any of three blind tests. Ratings for all brands made in a blind condition were not significantly different at the .05 level. The number of subjects able to identify their regular brand or any one of the four locally produced brands by taste in a blind test did not differ from chance. The number of subjects able to correctly match a milk sample on the basis of taste with another sample of the same brand from samples of the four brands did not differ from chance. In addition, subjects showed no ability to discern which sample was the highest in butterfat when one sample was more than 10 percent higher in butterfat than any of the other three samples.

Three main unexpected tendencies were found. First, a few subjects could discern Hansen's milk from all other samples in a blind test. Hansen's tended to be rated lower than the other three brands. Second, smokers tended to correctly identify their brand of milk in a blind test more
frequently than non-smokers. Smokers also tended to correctly match samples of milk in a blind test more frequently than non-smokers. Third, it was found that subjects who correctly matched milk samples rated their confidence in their decision significantly lower than subjects who incorrectly matched milk samples.

**Final Appraisal**

Further research is indicated to explain and to determine the extent of the ability of some subjects to identify Hansen's All Star brand milk as being different from other brands of 4 percent butterfat milk in a blind test. Research is needed to explain and to verify or to dismiss the observed tendency of smokers to identify and to match milk samples in blind tests with a greater frequency of correctness than non-smokers. Theory and research is needed to explain the lower confidence ratings correct subjects assigned to their decision in matching samples of milk. In addition, further research is needed to collaborate and to verify the lack of ability of subjects to discern among various brands of 4 percent butterfat milk.

Moreover, research should be performed to determine why consumers buy the brands of milk that they buy. It seems likely that physical differences between milk brands play only a minor role at best in consumer purchase decisions for 4 percent butterfat milk. The main marketing implication of this study suggests that factors other than product composition
should be stressed in marketing strategies. Perhaps, consumer perception of product attributes would prove to be the avenue to a larger market share since price is controlled by the state. Indeed, product differentiation seems to arise only through perceived attributes. Methods for altering and controlling consumer perceptions offer a large array of possibilities for future research.
REFERENCES

BOOKS


INTERVIEWS


PERIODICALS


Jacoby, Jacob; Olson, Jerry C.; and Haddock, Rafael A. "Price, Brand Name, and Product Composition Characteristics as Determinents of Perceived Quality." *Journal of Applied Psychology*, LV (1971), 570-579.


