



1993

## Effect of product branding, age and sex of rater, and type of judgement on presentation order bias in paired comparison taste testing

George Webster Hunt  
*Loyola University Chicago*

Follow this and additional works at: [https://ecommons.luc.edu/luc\\_diss](https://ecommons.luc.edu/luc_diss)



Part of the [Education Commons](#)

---

### Recommended Citation

Hunt, George Webster, "Effect of product branding, age and sex of rater, and type of judgement on presentation order bias in paired comparison taste testing" (1993). *Dissertations*. 3012.  
[https://ecommons.luc.edu/luc\\_diss/3012](https://ecommons.luc.edu/luc_diss/3012)

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact [ecommons@luc.edu](mailto:ecommons@luc.edu).



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](#).  
Copyright © 1993 George Webster Hunt

LOYOLA UNIVERSITY OF CHICAGO

EFFECT OF PRODUCT BRANDING, AGE AND SEX OF  
RATER, AND TYPE OF JUDGEMENT ON PRESENTATION  
ORDER BIAS IN PAIRED COMPARISON TASTE TESTING

A DISSERTATION SUBMITTED TO  
THE FACULTY OF THE DIVISION OF ARTS AND SCIENCES  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

DEPARTMENT OF  
COUNSELING AND EDUCATIONAL PSYCHOLOGY

BY  
GEORGE WEBSTER HUNT

CHICAGO, ILLINOIS  
JANUARY, 1993

Copyright by George Webster Hunt, 1992  
All Rights Reserved

## ACKNOWLEDGEMENTS

I thank my wonderful wife, Christina Hansen, for encouraging me to finish this work and giving me the emotional support I needed to make it happen. David Fishken helped to germinate the initial idea. Paul Riechmann provided an insightful critique of the proposal and the experimental design. Jan Fontanetta proofread the final document. I thank Bob Bailey for his support and for helping me find the time to do the work. Finally, I thank Tom Dutt and Peryam & Kroll for providing the data.

George Hunt is vice president associate research director at BBDO Chicago, an advertising agency in the BBDO Worldwide Network. He has been involved in work on sensory testing of products for development and claim substantiation since 1988. He received his B.A. in Engineering and Applied Science from Harvard College in 1979 and his M.Ed. in Community Counseling from Loyola of Chicago in 1984. His other work in measurement includes development of an instrument to measure perceived social support (Brown, S.D., Alpert, D., Lent, R.W., Hunt, G., & Brady, T. (1988). Perceived social support among college students: Factor structure of the Social Support Inventory. Journal of Counseling Psychology, 35(3), 472-478.). Mr. Hunt is a member of the American Marketing Association, the American Psychological Association, and the American Educational Research Association.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	iii
LIST OF TABLES .....	v
LIST OF FIGURES .....	vi
Chapter	
1. INTRODUCTION .....	1
2. REVIEW OF THE LITERATURE .....	10
3. METHOD .....	18
4. RESULTS .....	27
5. DISCUSSION .....	52
Appendix	
1. QUESTIONNAIRES .....	61
2. LEVEL OF "NO PREFERENCE" RESPONSE BY JUDGEMENT .....	67
REFERENCES .....	68

## LIST OF TABLES

Table	Page
1. Brand Preference by Order of Serving .....	3
2. Sample Age and Sex .....	19
3. Paired Comparison Questions .....	23
4. Classification of Paired Comparison Measures by Hedonic-type (H) versus Sensory-type (S) .....	26
5. Percent Choosing First Product by Blind vs. Branded Condition .....	29
6. Percent Choosing First Product by Sex .....	32
7. Percent Choosing First Product by Age .....	33
8. P-Values for Interactions among Branding, Sex, and Age .....	35
9. Ranking of 16 Paired Comparison Judgements on Proportion Choosing First Product .....	38
10. P-values associated with Differences in Proportion Choosing First Product between Sensory and Hedonic Judgements for 35 Sensory-Hedonic Judgement Pairs .....	39
11. Summary of Effects involving Product Branding that Vary across Sensory-Hedonic Judgement Pairs .....	41
12. Summary of Effects involving Respondent Sex that Vary across Sensory-Hedonic Judgement Pairs .....	44
13. Summary of Effects involving Respondent Age that Vary across Sensory-Hedonic Judgement Pairs .....	47
14. Correlations (Phi) between Sensory Judgements and Like Better Overall .....	56
15. Level of "No Preference" Response by Judgement .....	67

## LIST OF FIGURES

Figure	Page
1. Percent Choosing First Product by Age for 16 Paired Comparisons .....	34
2. Branding by Sex Interaction on Like Better Overall .....	36
3. Age Effect among Males for Like Better Overall vs. Longer Lasting Taste .....	49
4. Age Effect among Females Tasting Blind Products for Better Aroma vs. Longer Lasting Taste .....	50
5. Percent Choosing First Product by Age for Sticks More vs. Better Aroma and Better Flavor .....	51

## CHAPTER 1

### INTRODUCTION

The method of paired comparisons is widely used in marketing research to study consumer preference with respect to two or more consumer products, particularly food products. In this method, consumers taste two food products in succession separated by a short time interval during which the palate is typically cleared by eating a cracker and/or drinking a glass of water. After tasting both products, consumers are asked to indicate which of the two they prefer. Consumers may also be asked to make additional comparative judgements (e.g., which product is sweeter, which has the better texture, etc.). Results of these tests are used to make marketing decisions. For example, a food manufacturer might have consumers compare its new product to an existing competitive product. Test results might then be used as input into a decision of whether or not to launch the new product. Alternatively, paired comparison testing might be used to substantiate an advertising claim (e.g., "In blind taste tests, cola drinkers prefer the taste of Pepsi over Coke by a two-to-one margin"). It should be noted that use of the paired comparison methodology to substantiate advertising claims regarding consumer preference is particularly prevalent because this type of methodology is required by many of the bodies that currently regulate the use of comparative claims in advertising (Buchanan & Smithies, 1991).

A widely-recognized problem associated with the use of paired comparison taste testing is presentation order bias. This bias refers to

a tendency of subjects' preferences to be influenced by the order in which the products are sampled. For the most part, it has been found that respondents favor the first product sampled (primacy bias), but in some cases they favor the second product (recency bias). As a consequence of this bias, in a standard taste test it is recommended that the order in which the products are tasted is balanced across subjects. Taking the example introduced above, let's say Pepsi is preferred over Coke by 66% of the sample. When the results for the above example are analyzed by order of serving (see Table 1), a typical finding would be that the majority of consumers who try Coke first prefer Coke (say, 59% choose Coke, 41% choose Pepsi) and the majority of those who try Pepsi first prefer Pepsi (73% choose Pepsi, 27% choose Coke). If the order in which the products were served in this study had not been balanced across respondents, the results would be biased toward the product appearing more frequently in the first position. Further, even given the precaution of balancing the order of presentation, the presence of such order bias may reduce the face validity of the overall results.

Published discussions of order bias in consumer product tests first appeared in the late 1950s. Greenberg (1958) reported a study of two varieties of women's stockings tested by 516 female factory employees. The women wore the first pair of stockings until they wore out and then rated them on a number of dimensions. The second pair was similarly worn until worn out and then rated. Each variety was worn first by half of the sample. Ratings were found to be more positive on most dimensions for the stockings worn first.

Food scientists have noted the existence of order bias in the rating of multiple food samples. Eindhoven, Peryam, Heiligman, and Hamman (1964)

Table 1

Brand Preference by Order of Serving


---

		Served	Served
	<u>Total</u>	<u>Pepsi first</u>	<u>Coke first</u>
<u>Preference</u>			
Pepsi	57%	73%	41%
Coke	<u>43%</u>	<u>27%</u>	<u>59%</u>
	100%	100%	100%

---

found a position effect in ratings of meat. They had subjects taste four meat products, rating each on a nine-point hedonic scale. Samples tried earlier in the series were found to be rated higher. Kamen, Peryam, Peryam, and Kroll (1970) had subjects rate from 4 to 12 food samples selected from four different types of food and found that hedonic ratings of the samples were consistently lower for samples tasted in the latter half of the session. While these studies did not use paired comparison methodology, they illustrate the general phenomenon of order effects in taste tests.

Like food scientists, marketing researchers have also found order bias in tests of food products. In a paired comparison test of two types of matzo bread, Greenberg (1963) found that while there was not a significant overall preference for either product, each product was preferred over the other among consumers tasting it first. Berdy (1969) aggregated findings from six paired comparison taste tests involving cakes and breakfast cereal and reported that

two-thirds of respondents chose the first product tried. First product paired comparison bias has also been reported for soups (Scowcraft, 1958), and for fried foods, and soft drinks (Dean, 1980).

It should be pointed out that order effects are not always present. Greenberg (1958) reported no order bias in the stockings study on a paired comparison question ("Which pair would you rather buy?"). Day (1969) reported finding no order bias in 15 paired comparison tests involving all possible pairs of six experimental food products.

Even when there is evidence of order bias in paired comparison testing, it does not invariably favor the first product served. Penny, Hunt, and Twyman (1972) examined 247 food paired comparison tests and found that the first product was favored by at least two percentage points in 60 cases (24%), there was no difference in 72 cases (29%), and the second product was favored by at least two points in 115 cases (46%). When they included products from non-food categories (household products and toiletries), the 463 total cases were found to be almost evenly distributed across the three outcomes. The first product was favored in 145 cases (31%), no difference was found in 153 cases (33%), and the second product was favored in 165 cases (36%). It is difficult to evaluate the robustness of the Penny et al. findings because they did not report the sample sizes involved in their tests. However, their test-wise results clearly show that order effects in taste tests may not be limited to primacy effects.

If order effects don't always occur, what factors are associated with their presence? Eindhoven et al. (1964) suggested that order effects are not inevitable, but are affected by various aspects of the experimental situation, including food type and subject expectations. Day (1969) hypothesized that

order effects are most likely to occur when differences between products being compared are small. He suggested that the absence of order effects in his study may have been because the members of each pair were chosen to differ widely on a set of sensory characteristics, in contrast to the relative similarity of the marketplace products typically tested. He suggested that when consumers have difficulty discriminating between the products being compared, cues that might not otherwise have an effect (such as the novelty of the first product) may influence decision-making.

To test this hypothesis, Day conducted a small sample study in which he ran two paired comparison tests. The first pair of products had obvious differences in texture and flavor, and the second pair was actually two samples of the same product. Day found no order effect with the first pair and a directional but non-significant effect with the second pair. It should be noted that a limitation of Day's study was the small sample sizes he employed ( $n=85$  for each test).

Mitchell (1956) performed a similar experiment in which he tested four pairs of whiskey samples varying within pair from identical to very different (based on discrimination testing). Mitchell found a significant bias in favor of the first product when aggregating across the four pairs. He also found that the tendency to favor the first product increased as product similarity increased, from a non-significant bias for the most different pair to a highly significant bias for the identical pair.

Mitchell also attempted to explain what causes preference for the first product when discrimination is difficult. He hypothesized that the second stimulus is less pleasurable than the first because it reinforces the "neurological

trace" laid down by the first stimulus. He reasoned that the second of two taste stimuli was perceived as being more intense and could therefore be perceived as being less pleasant. Mitchell did an experiment to test the trace theory by conducting a paired comparison test with two whiskeys in which half the sample tasted a non-test "conditioner sample" before trying the pair of test products. He found a primacy effect in the group without the conditioner sample and no order effect in the group that was exposed to the conditioner sample. He concluded that this finding was consistent with the trace theory.

One problem with Mitchell's trace explanation is that it necessitates placing an upper limit on the aversiveness of reinforcing stimuli. It could be argued that although the first stimulus becomes aversive when it follows a conditioner sample, the second stimulus is even more aversive since it now follows two like stimuli. If this were to occur, the primacy bias would exist even when the test was preceded by a conditioner sample.

As noted above, it is standard practice to rotate the order of stimulus presentation to neutralize the effect of order on the preference measure. However, neutralizing the effect does not eliminate it. Clearly, even balanced order effects are a nuisance in that they may cloud measurement of true preference. An increased understanding of factors influencing order effects could have practical application in the design of product tests and theoretical significance with respect to the knowledge base related to how subjects respond to paired comparison tasks. Given the widespread usage of the method of paired comparisons, such an understanding seems well worth pursuing.

One aspect of order bias in the method of paired comparisons that has not been reported in the literature is the effect of product branding on the presence and size of order effects. The majority of taste test work is conducted on a "blind" basis with the consumer being unaware of the brand names of the products compared. In many cases, however, companies are interested in the preferences between two products when the brand names are part of the stimuli. Following the line of reasoning advanced earlier, it is hypothesized that branding will be associated with lower task difficulty due to increased information about the stimuli. In this instance less order bias would be expected.

A second aspect of order bias that has not been systematically studied is the relationship between respondent characteristics and order bias. Two such characteristics are age and sex. Younger subjects are known to have greater taste abilities in some situations than older subjects and there is evidence that females have taste abilities superior to males in some circumstances (Coward, 1989; Wysocki & Gilbert, 1989). Again, extending on Day's theory, one would expect younger subjects to be less prone to order bias than older subjects and females to be less prone than males as these groups are likely to find the discrimination task less difficult and therefore be less likely to rely on information external to the products being tested in making judgements.

A third area that has not been reported on in the literature is the degree to which order bias depends on the type of judgement being made. The studies discussed above focussed principally on hedonic judgements (i.e., judgements of which of two products is liked most). Another name for such judgements would be evaluative judgements. Many paired comparison taste tests also

include sensory judgements (i.e., judgements of which of two products is sweetest, softest, stickiest, etc.). It can be argued that making hedonic judgements is more difficult than making sensory judgements as the former may require integrating several sensory dimensions to arrive at an overall judgement. Extending Day's theory that discrimination difficulty leads to greater order effects, less order bias would be expected on sensory judgements than on hedonic judgements.

The overall purpose of the investigation to be reported here is to expand the knowledge base with regard to specific factors (product branding, respondent age and sex, and type of comparison) influencing presentation order bias in paired comparison taste tests. Specifically, an attempt is made to expand on the work of Day (1969) and Mitchell (1956) who reported that order effects increase when respondents have difficulty discriminating between the products being compared. A data set from a large-scale paired comparison study (n=8000) of ten chewing gums was systematically examined to test for the possible influence of four factors (presence/absence of branding, respondent age, respondent sex, and type of comparison) on the presence and strength of order effects. An order effect in this study is defined to exist at the group level if the first product sampled is chosen more often than the 50% expected by chance. Strength of order effect is defined as amount of deviation from 50%.

### Research questions

Five research questions are addressed. First, is the occurrence and strength of order bias affected by the presence/absence of branding? Second, is the occurrence and strength of order bias affected by the sex of the

respondent. Third, is the occurrence and strength of order bias affected by the age of the respondent? Fourth, is the occurrence and strength of order bias affected by the type of comparative judgement being made? Finally, do the effects of branding and respondent age and sex interact with judgement type?

Five research hypotheses are tested:

- H<sub>1</sub>. Order effects are stronger for unbranded product pairs than for branded product pairs.
- H<sub>2</sub>. Order effects are stronger for male respondents than for female respondents.
- H<sub>3</sub>. Order effects are stronger for older respondents than for younger respondents.
- H<sub>4</sub>. Order effects are stronger with hedonic-type judgements (e.g., "like better overall", "better flavor") than with sensory-type judgements (e.g., "softer", "sweeter", "stronger").
- H<sub>5</sub>. Product branding and respondent age and sex interact with judgement-type such that their effects on order bias are strongest for hedonic-type judgements and weakest for sensory-type judgements.

## CHAPTER 2

### REVIEW OF THE LITERATURE

In Chapter 1, a selected review of the research directly bearing on presentation order bias in paired comparison taste tests was presented. In this chapter, a general overview of applications of paired comparison methodology in psychological research is provided, along with a discussion of its role in consumer product research. Finally, a brief review of taste test research related to the variables studied in this investigation is presented. In terms of the dependent variable, a brief review of order effects in other areas of research is provided. For each independent variable (product branding, respondent age and sex, and hedonic vs. sensory judgement-type), findings from available taste research are reviewed. It should be pointed out that because much of consumer taste test research is conducted commercially, the number of published studies in most of these areas is limited.

#### The paired comparison task in psychological research

The paired comparison task has a long history of use in psychological research. Over a 100 years ago, Fechner (1860) pioneered the technique in weight-estimation experiments in which he would first lift a standard weight and then judge the weight of a second object relative to the standard. Thurstone (1927) later popularized the use of multiple paired comparisons to construct attitude scales. It continues today to receive heavy use in scaling research (van

der Ven, 1980).

The method of paired comparisons is currently used in the study of a wide variety of problems including taste testing, sensory perception (Schneider, 1980), infant cognition (Fagan, 1981; MacKay-Soroka, Trehub, Bull, & Corter, 1982), interest inventories (Edwards, 1959), personnel ratings (Lawshe, Kephart, & McCormick, 1949), teacher evaluation (Ory, 1980), jury verdicts (Tetlock, 1983), and the general study of choice behavior (e.g., Luce, 1959, 1977).

There is also a substantial literature on the design of paired comparison experiments, stochastic modeling of paired comparison situations, and analysis of paired comparison data. The interested reader is referred to reviews by Bradley (1976), Davidson and Farquhar (1976), and David (1963).

#### Use of paired comparison task in consumer taste tests

Paired comparison methodology is widely used in consumer taste testing. It is one of four types of procedures identified by Batsell and Wind (1979) in their review of current product testing methods. The other methods are monadic (single product) tests, staggered tests (multiple product tests without direct comparisons), and conjoint analysis (multiple product tests with products systematically designed to vary on one or more factors). Moskowitz (1983), a leading product testing consultant, identified paired comparison testing as "one of the more ubiquitous measuring procedures in the testing armory". An often-cited reason for the prevalence of the technique is its strong face validity, particularly with non-technical top executives at large food companies.

### Research on order effects in paired comparison testing

Order effects on response to sequentially presented stimuli have been observed in several areas of psychological research. Primacy and recency effects in memory research (i.e., the phenomenon of items learned first and last being better recalled) have been demonstrated in countless studies. The subject has also received attention from social psychologists studying the effect of order of argument presentation on persuasion (e.g., Lana, 1961) and on impression formation (Dreben, Fiske, & Hastie, 1979).

Closer to taste testing is the work that has been done on stimulus presentation order effects in sensory perception research. Allan (1984) had subjects judge which of two brief auditory tones was longer in duration and found a bias toward the first tone. Jamieson and Petrusic (1976) found that subjects tended to choose the second tone more frequently, but that the direction of the order bias was reversed when interstimulus time was increased and was eliminated when feedback concerning task performance was provided to subjects.

Presentation order bias has also been reported on tasks involving other sensory modalities. Mitchell (1956) cited research showing order bias in judgements of visual stimuli (lines, circles, and squares), auditory stimuli (musical chords), and odors. The presence and the direction of the bias, he reported, varied with the type and range of stimuli, the time interval between stimuli, the type of judgement, subject training, and other criteria.

### Effect of product branding on taste test outcomes

The majority of taste tests are conducted "blind" without identifying the brand name or product claims. There are situations, however, when marketers are interested in comparing results of blind testing with results from the same test conducted with branded products. Because of expectations set up by advertising and prior experience with a brand, results of branded testing can differ from blind testing. One possible situation calling for branded testing would be when a new product is a winner over a leading competitor in blind testing and the manufacturer wants to determine whether or not the new product will be able to overcome the equity of positive attitudes toward the established brand.

There is no published research on the effect of product branding on order effects in paired comparison taste tests or on any effects related to branding in paired comparison tests. The available research on the effect of branding has been done in sequential monadic tests. Arndt (1970) had consumers rate beers in blind and branded conditions and found ratings for five of six products to be higher branded than blind. Moskowitz (1985) studied several brands of coffee in both blind and branded conditions. He found that users of a brand tended to upgrade their ratings of that brand when in the branded condition, but that non-users of the brand did not rate it differently branded than blind. Moskowitz also observed that the more subjective attributes (such as flavor quality) were affected more by branding than were the less subjective attributes (such as darkness of granules).

### Age and sex effects in sensory research

It is generally accepted that people can perceive different things from a common stimulus. The effect of age and to a lesser extent sex on sensory perception have been studied. Moskowitz (1985) reported that while studies on the senses of vision and audition suggest that children and adults do not differ in their response to tones and lights, we know less about age differences with regard to the chemical senses of taste and smell.

Measurement of taste and smell behavior is dependent on a number of factors, including how the taste/smell function is characterized and the stimuli. Wysocki and Gilbert (1989) discussed four ways to characterize functioning: (a) Threshold; (b) Magnitude estimation (strength or hedonic quality); (c) Similarity judgements; and (d) Identification. Their study focussed primarily on identification of odors. They also collected self-rating of olfactory ability. They found that women had better olfactory ability than men on both identification and self-rating. They also found a general decline in olfactory ability with age, but usually not beginning until the fifth, sixth, or in some cases the eighth decade, depending on the odor tested.

Cowart (1989) studied gustatory and olfactory ability of 137 adults ranging in age from 19 to 79. She found that females had a slightly lower taste threshold than males for sodium chloride and citric acid. She found no difference in threshold by age. On taste identification, however, she found significant age effects (younger outperforming older) in addition to gender effects (females outperforming males). It should be noted that this study did not include non-adults and the sample size was not sufficient to determine whether

ability declines uniformly with advancing age or only after middle adulthood, as found in the Wysocki and Gilbert study.

Research on the effects of respondent age and sex on paired comparison preferences is not extensive. Buchanan (1987, 1988) has developed models and conducted empirical investigations of respondent discrimination ability in paired comparison testing. Buchanan has recognized that although respondents are forced to choose between the two products in a standard paired comparison test, discrimination between the two products (a necessary condition for true preference) is a probabilistic process. Consequently, among respondents choosing a given product, some have truly discriminated between the alternatives and others have selected it randomly. Using techniques such as repeat paired comparisons or "triangle testing" (i.e., picking the odd product from a set of three products of which two are identical) individual and group discrimination ability can be estimated.

Buchanan, Givon, and Goldman (1987) recruited 180 subjects in a shopping mall and had them try two cola formulations. Each subject performed four paired comparison tests and a four-trial triangle test. On ability to discriminate, heavy users were better than light users, but there were no differences between males and females or across age groups. Two limitations of this study are that it was conducted on a single product category and that the age range (not reported) may not have been extensive.

There is some evidence that younger respondents have a tendency to express less extreme preferences than older respondents. Engen (1974) had children and adults rate a set of odors in a series of paired comparisons and used Thurstone scaling to construct hedonic scales for the two age groups. He

found that the liking scale on the odors tested had a considerably smaller range for children than for adults. K. Kraska (personal communication, March 23rd, 1992) reported that teens typically showed smaller preference skews than adults in her work with paired comparison testing on beverages, cereals (both ready-to-eat and hot) and candy. It is important to note that these investigators have not established whether younger respondents actually experience less difference between stimuli or merely express less consistent preferences on paired comparison tasks.

In summary, the available research on the effects of age and sex on the sense of taste is not extensive. Available evidence suggests that ability may be greater for younger respondents and for females. However, the effects appear dependent on the exact stimulus being rated and on the task demands.

### Hedonic versus sensory judgements

Typologies of perception into categories labeled cognitive ("I get it") and affective ("I like it") are common in psychology. In taste test research, the primary focus is on the latter type of judgement, typically termed "hedonic" response. Taste testing is also concerned with the other type of judgement, known variously as "descriptive", "objective", or "pure sensory". Practically, sensory-type judgements can be useful in learning what drives hedonic judgements.

It does appear that hedonic and sensory measures are tapping into different domains. A study of taste discrimination and hedonic response to sucrose in coffee conducted by Lundgren (1978) found that ability to discriminate among different sucrose levels and degree of liking for sucrose

levels in coffee are independent behavioral responses. In the Moskowitz (1985) study of blind versus branded coffee ratings, objective measures were found to be less affected by product branding than hedonic measures.

Logically, however, the two domains are not completely independent. It is likely that hedonic responses are based on characteristics that can be objectively described. Also, as discussed earlier, ability to discriminate between two products is a necessary condition for preferring one of them.

In summary, the method of paired comparisons is used to study a wide range of phenomena. It is one of the major tools used in consumer taste testing. Order effects associated with paired comparisons have been studied in a number of fields. As discussed in Chapter 1, order effects in taste tests typically favor the first product and tend to be greater when the products compared are very similar. Product branding in consumer taste tests has been found to raise product ratings in some circumstances, but its effect on paired comparison data has not been studied. With regard to respondent characteristics, females and younger respondents tend to have higher ability and younger respondents tend to show a restricted range on paired comparison preference. Finally, there is some evidence that hedonic and sensory measures are tapping into different domains.

## CHAPTER 3

### METHOD

The data were obtained from a large-scale commercial study of consumer preferences relating to ten chewing gum products conducted in 1989.

#### Subjects

Subjects were 8000 male and female consumers between the ages of 12 and 55 recruited at ten shopping malls geographically dispersed across the U.S.

#### Procedure

Subjects were intercepted in shopping malls and systematically screened with respect to age, sex, race, and past week gum usage. Age within sex and race quotas were set so that the total sample was representative of past week users of chewing gum. Sample sizes by age within sex are shown in Table 2. Subjects were told they would be asked to evaluate two chewing gums during a 30 minute time period and that they would be given \$3.00 for participating. Subjects agreeing to participate were escorted to a room and seated at a table.

The subjects were randomly assigned to one of 36 conditions. These conditions consisted of 18 product pairs crossbroken with respect to two levels of branding (branded and unbranded). Half of the subjects were given branded

products; the other half unbranded products. The samples were number coded and order of serving was balanced across subjects so that each product appeared in first position an equal number of times. Order of serving was also balanced as much as possible for age within sex for each of the 36 conditions.

Each subject was given a single serving of the first sample on a paper plate. In the branded condition, the product was wrapped in its standard single-serving packaging which displayed the brand name of the product being tested. In the blind condition, the product was unwrapped by the monitor prior to serving. The wrapper was kept out of the subject's sight. Subjects were

Table 2

Sample Age and Sex

	<u>Female</u>	<u>Male</u>	<u>Total</u>
12-17	674 (8.4%)	464 (5.8%)	1138 (14.2%)
18-34	2304 (28.8%)	1559 (19.5%)	3863 (48.3%)
35-55	<u>1731 (21.6%)</u>	<u>1268 (15.9%)</u>	<u>2999 (37.5%)</u>
Total	4709 (58.9%)	3291 (41.1%)	8000 (100.0%)

instructed to chew the first gum for 12 minutes and to complete product ratings on a self-administered questionnaire at first bite and after 1, 3, 6, 9, and 12 minutes. Subjects used oven timers to determine when each set of ratings should be completed. At the conclusion of the first product trial, subjects

were instructed to drink a glass of mineral water and after six minutes were given the second sample.

Immediately following the 12 minute chew of the second product, subjects were instructed to complete a questionnaire requiring them to make 16 judgements related to the two products just sampled. For overall preference, the first paired comparison question, subjects were forced to make a choice (i.e., a "no preference" or tie judgement was not allowed). However, on the 15 subsequent paired comparison judgements, a "no preference" decision was accepted if volunteered. For each subject, the order of serving for the product chosen on each of the 16 paired comparison judgements was recorded.

### Instrumentation

The instrumentation consisted of two paper-and-pencil questionnaires, a product rating questionnaire and a paired comparison questionnaire. The product rating questionnaire was completed for each gum during the 12 minutes it was chewed. The paired comparison questionnaire was filled out after the 12 minutes of chewing had been completed for both gums. A monitor blind to the purposes of the study was present at all times to answer questions and read instructions where necessary.

The product rating questionnaire is not analyzed in the current study, but is described here for completeness. It consisted of closed-ended ratings of the gum on a number of characteristics including flavor quality, texture, and sticking to teeth. Subjects were instructed to rate the gum at first bite and then at 1, 3, 6, 9, and 12 minutes. Two types of scales were used primarily: a five-point excellent to poor scale (e.g., Flavor: Excellent, Good, Average, Poor, Very

Poor) and a three-point satisfaction scale (e.g., Texture: Too soft, About right, Too hard). Subjects were also asked to rate each gum on a nine-point hedonic scale and a five-point purchase intent scale, and to record open-ended likes and dislikes.

The paired comparison questionnaire was completed immediately after the subject finished the product rating questionnaire for the second gum. The paired comparison questionnaire consisted of 16 paired comparison questions and one open-ended question. The first question was overall preference ("Which gum did you like better, everything considered?"). Subjects were asked to write in the code number (blind condition) or brand (branded condition) of the brand they liked best. Subjects were then asked to write down why they liked the gum they chose better. Next, subjects were asked to respond to 15 additional paired comparisons (e.g. "Which gum was softer?", "Which gum had the better flavor?"). The particular comparisons included were chosen to address marketing questions about the brands tested. The full list of paired comparison questions are shown in Table 3. A complete set of questionnaires is in appendix 1.

### Hypotheses

The following null hypotheses were tested:

- Ho<sub>1</sub>. There is no difference in strength of order effects across unbranded and branded product pairs.
- Ho<sub>2</sub>. There is no difference in strength of order effects between males and females.
- Ho<sub>3</sub>. There is no difference in strength of order effects across age groups.

- H04. There is no difference in strength of order effects across hedonic-type judgements (e.g., "like better overall", "better flavor") and sensory-type judgements (e.g., "softer", "stronger", "cooled mouth more").
- H05. There is no interaction between type of judgement (hedonic versus sensory) and product branding, type of judgement and sex, and type of judgement and age.

Table 3

Paired Comparison Questions

---

1. Which gum did you like better, everything considered?
  2. Which gum had the better flavor?
  3. Which gum had the better texture?
  4. Which gum had the longer lasting taste?
  5. Which gum was more refreshing?
  6. Which gum was softer?
  7. Which gum was sweeter?
  8. Which gum had the better aroma (smell)?
  9. Which gum had the more refreshing flavor?
  10. Which gum had the stronger flavor?
  11. Which gum cooled your mouth more?
  12. Which gum made your mouth feel fresher?
  13. Which gum sticks more to your teeth or dental work?
  14. Which gum moistens your mouth better?
  15. Which gum was better for freshening the breath?
  16. Which gum would you rather buy?
-

### Design and analysis

Each subject was assigned to one of two groups for each of the 16 paired comparison judgements (chose the first product served or chose the second product served). The resulting data were analyzed using multinomial ANOVA (Woodward, Bonett, and Brecht, 1990). It should be noted that this analytic technique, unlike ANOVA proper, does not require that the dependent variable be normally distributed. The only assumptions required are that the responses of each subject can be classified into one and only one category and that the responses of one subject do not affect the responses of any other subject. Hypotheses are tested using the Wald statistic based on the multinomial distribution.

The basic design is a factorial 2x2x3 design, with two levels of branding (blind/branded), two gender levels, and three age levels (12-17, 18-34, 35-55). There are 16 dependent variables consisting of the 16 paired comparison judgements.

The first three null hypotheses are tested by examining the main effects for product branding, respondent sex, and respondent age for each of the 16 judgements.

The fourth null hypothesis is tested via a contrast between the hedonic-type and sensory-type judgements for the total sample. Table 4 shows how the 16 judgements are classified into hedonic-type vs. sensory-type. It should be noted that 4 of the 16 judgements have not been classified because they do not clearly belong to one of the two categories. Order effects for the five hedonic-type judgements (Like better overall, Better flavor, Better texture, Better aroma, and Rather buy) are contrasted with order effects for the seven sensory-

type judgements (Longer lasting taste, Softer, Sweeter, Stronger flavor, Cooled mouth more, Made mouth feel fresher, and Sticks more to teeth/dental work).

The final null hypothesis is tested using a series of 35 MANOVA analyses involving all possible pairs of hedonic and sensory judgements (5 hedonic X 7 sensory = 35). For each analysis, the effects of product branding, respondent sex, and respondent age and their interactions on the difference in order effect between the hedonic and sensory judgements are examined using a test for marginal homogeneity across the dependent variables.

Table 4

Classification of Paired Comparison Measures by Hedonic-type (H)  
versus Sensory-type (S)

---

1. Like better overall (H)
  2. Better flavor (H)
  3. Better texture (H)
  4. Longer lasting taste (S)
  5. More refreshing
  6. Softer (S)
  7. Sweeter (S)
  8. Better aroma (H)
  9. More refreshing flavor
  10. Stronger flavor (S)
  11. Cooled mouth more (S)
  12. Made mouth feel fresher (S)
  13. Sticks more to teeth or dental work (S)
  14. Moistens mouth better
  15. Better for freshening the breath
  16. Rather buy (H)
-

## CHAPTER 4

### RESULTS

#### Hypothesis 1

The first hypothesis stated in the null form (There is no difference in strength of order effects across unbranded and branded product pairs.) was rejected for overall preference and for the majority of the other paired comparisons. The proportion of subjects choosing the first product on Like better overall is 5.2 percentage points higher in the blind condition (63.4%) than in the branded condition (58.2%).

Table 5 shows the percentage of subjects choosing the first product tried across blind and branded conditions for each of the 16 paired comparison judgements. The judgements are displayed in order of blind versus branded difference from largest to smallest difference. (It should be noted that for each judgement percentages are based on respondents expressing a preference. "No preference" responses or abstentions were infrequent for most judgements (1-3%). Sticks more to teeth/dental work was an exception with a 46.9% non-response. Complete data on non-response by judgement is shown in Table 15 in appendix 2.)

All judgements, with the exception of Sticks more to teeth, showed a first product bias in both blind and branded conditions. Subjects rating products in the blind condition were significantly more likely than those in the branded condition to choose the first product served on 11 of 16 judgements. The size of

of the branding effect ranged from 5.5 points for Rather buy to 2.8 points for Longer lasting taste.

Judgement of which product sticks more to teeth or dental work showed a bias toward the second product. The order bias was again stronger for subjects in the blind condition, with 57.4% choosing the second product versus 53.5% for subjects evaluating branded products. Again, It should be noted that close to half of the sample indicated that they had no preference on this measure.

Three judgements showed no order effect due to branding. Sweeter, Softer, and Stronger flavor showed a slight tendency toward first product bias in both blind and branded conditions with about 54% of the sample choosing the first product on each judgement. It is important to note here that all three of these judgements are sensory judgements and that the sensory judgements generally show smaller effects due to branding than the hedonic judgements. This finding will be explored in greater detail in the results section relating to testing hypothesis 5.

### Hypothesis 2

The second hypothesis stated in the null form (There will be no difference in strength of order effects between males and females.) was rejected for overall preference and for the majority of the other paired comparisons. However, the direction of the difference was found to be opposite to that predicted. Females were more likely than males to choose the first product, with 62.6% choosing the first product on Like better overall versus 58.1% for males.

Table 6 shows the percentage of subjects choosing the first product tried for each of the 16 paired comparison judgements by sex. The judgements are

Table 5

Percent Choosing First Product by Blind vs. Branded Condition


---

	<u>Blind</u>	<u>Branded</u>	<u>Difference (p-value)</u>	
	(n=4000)	(n=4000)		
Rather buy (H)	64.0%	58.5%	5.5%	(.001)
Like better overall (H)	63.4	58.2	5.2%	(.001)
Better flavor (H)	62.4	57.8	4.6%	(.001)
Better texture (H)	61.8	57.7	4.1%	(.007)
More refreshing flavor	62.8	58.7	4.1%	(.001)
Better for freshening the breath	62.6	58.5	4.1%	(.001)
More refreshing	63.1	59.4	3.7%	(.001)
Better aroma (H)	66.0	62.3	3.7%	(.003)
Cooled mouth more (S)	62.4	58.9	3.5%	(.001)
Made mouth feel fresher (S)	62.2	58.7	3.5%	(.004)
Moistens mouth better	57.7	54.7	3.0%	(.049)
Longer lasting taste (S)	61.8	59.0	2.8%	(.013)
Sweeter (S)	54.6	53.0	1.6%	(>.500)
Softer (S)	54.3	53.7	0.6%	(.198)
Stronger flavor (S)	55.7	55.9	-0.2%	(>.500)
Sticks more to teeth or dental work (S)	42.6	46.5	-3.9%	(.011)

---

Note. n slightly smaller on judgements other than "Like better overall" due to small number of "no preference" responses.

displayed in order of difference between males and females from the largest to smallest difference. All judgements with the exception of Sticks more to teeth showed a first product bias among both males and females. Females, however, were found to be significantly more likely than males to choose the first product served on 13 of 16 judgements. The difference in order bias between males and females ranged from 5.5 points for Cooled mouth more to 2.3 points for Stronger flavor. The effect sizes are comparable to those found for branding.

Sweeter, Softer, and Sticks more to teeth/dental work showed no significant differences between males and females. As was found with the branding effect, the judgements on which no significant effect were found were sensory judgements. There were, however, several sensory judgements with large sex effects (e.g., Cooled mouth more). The interaction of sex and judgement-type is explored in greater detail in the results section relating to testing hypothesis 5.

### Hypothesis 3

The third hypothesis stated in the null form (There will be no difference in strength of order effects across age groups.) was not rejected for overall preference. There was not a significant difference in the proportion of subjects selecting the first product on Like better overall across the three age groups (12-17--58.0%; 18-34--60.9%; 35-55--61.7%;  $p=.112$ ).

There were significant age effects, however, on 7 of the 16 judgements. These judgements included both sensory and hedonic judgements: More refreshing, Stronger flavor (S), Better flavor (H), Sweeter (S), Moistens mouth better, More refreshing flavor, and Longer lasting taste (S). Several of these judgements relate in some way to the intensity of the taste.

Table 7 shows the percentage of subjects choosing the first product tried for each of the 16 paired comparison judgements by age. The judgements are displayed in order of effect size from largest to smallest. For the seven judgements with significant age effects, the differences are driven primarily by the 12-17 age group. Figure 1 shows that the 12-17 year olds show a lower tendency to choose the first product than the other two age groups.

Interactions among product branding, sex of respondent, and age of respondent were examined for each of the 16 paired comparisons. The p-values are shown in Table 8. One significant ( $p < .05$ ) interaction was found: a branding by sex interaction on the Like better overall judgement. The interaction is ordinal. As shown in Figure 2, females in the blind condition have an even larger first order bias than would be predicted by summing the main effects of branding and sex.

#### Hypothesis 4

The fourth hypothesis stated in the null form (There will be no difference in strength of order effects across hedonic-type judgements and sensory-type judgements.) was not rejected. Order effects for a subset of the sensory judgements were found to be significantly weaker than for the hedonic judgements, but several of the sensory judgements exhibited order effects as strong as the hedonic judgements.

Table 9 shows the 16 paired comparisons ranked by percent choosing product served first among the total sample. All judgements show significant order effects. Descriptively, there are three tiers of order effects. Judgements in the first tier, consisting of the 11 highest-ranking judgements, show order effects

Table 6

Percent Choosing First Product by Sex


---

	<u>Females</u>	<u>Males</u>	<u>Difference (p-value)</u>	
	(n=4709)	(n=3291)		
Cooled mouth more (S)	62.9%	57.4%	5.5%	(.001)
More refreshing	63.4	58.0	5.4	(.001)
Better aroma (H)	66.3	61.0	5.3	(.001)
Made mouth feel fresher (S)	62.6	57.5	5.1	(.001)
Moistens mouth better	58.2	53.3	4.9	(.004)
More refreshing flavor	62.7	57.9	4.8	(.001)
Rather buy (H)	63.2	58.5	4.7	(.001)
Like better overall (H)	62.6	58.1	4.5	(.002)
Better for freshening the breath	62.4	57.9	4.5	(.001)
Better flavor (H)	61.9	57.5	4.4	(.001)
Longer lasting taste (S)	61.8	58.4	3.4	(.002)
Better texture (H)	61.0	58.0	3.0	(.028)
Stronger flavor (S)	56.7	54.4	2.3	(.019)
Sweeter (S)	54.0	53.5	0.5	(.479)
Softer (S)	54.0	54.0	0.0	(>.500)
Sticks more to teeth/dental work (S)	43.8	45.3	-1.5	(.368)

---

Note. n slightly smaller on judgements other than "Like better overall" due to small number of "no preference" responses.

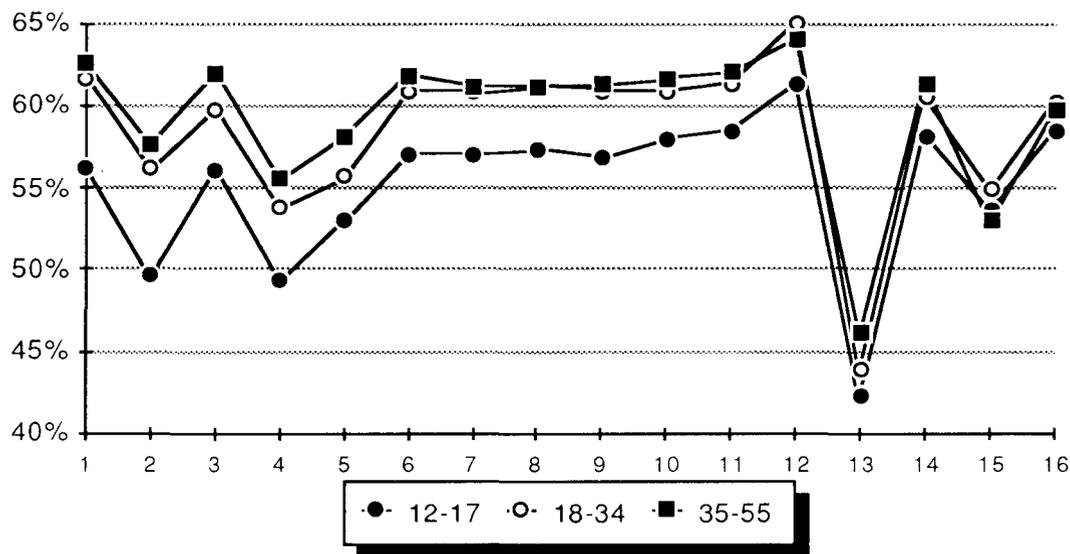
Table 7

Percent Choosing First Product by Age

	<u>12-17</u>	<u>18-34</u>	<u>35-55</u>	<u>P-value</u>
	(n=1138)	(n=3863)	(n=2999)	
More refreshing	56.2%	61.7%	62.6%	.001
Stronger flavor (S)	49.6	56.2	57.7	.001
Better flavor (H)	56.1	59.8	62.0	.002
Sweeter (S)	49.3	53.8	55.5	.002
Moistens mouth better	53.0	55.7	58.2	.019
More refreshing flavor	57.1	60.9	61.9	.023
Longer lasting taste (S)	57.1	60.8	61.2	.041
Cooled mouth more (S)	57.4	61.2	61.2	.054
Better for freshening the breath	56.8	60.9	61.3	.109
Like better overall (H)	58.0	60.9	61.7	.112
Rather buy (H)	58.5	61.4	62.1	.125
Better aroma (H)	61.3	65.0	64.1	.165
Sticks more to teeth (S)	42.3	43.9	46.2	.179
Made mouth feel fresher (S)	58.2	60.5	61.3	.229
Softer (S)	53.7	54.9	53.0	.339
Better texture (H)	58.4	60.2	59.7	>.500

Note. n slightly smaller on judgements other than "Like better overall" due to small number of "no preference" responses.

Figure 1

Percent Choosing First Product by Age for 16 Paired Comparisons

## Key:

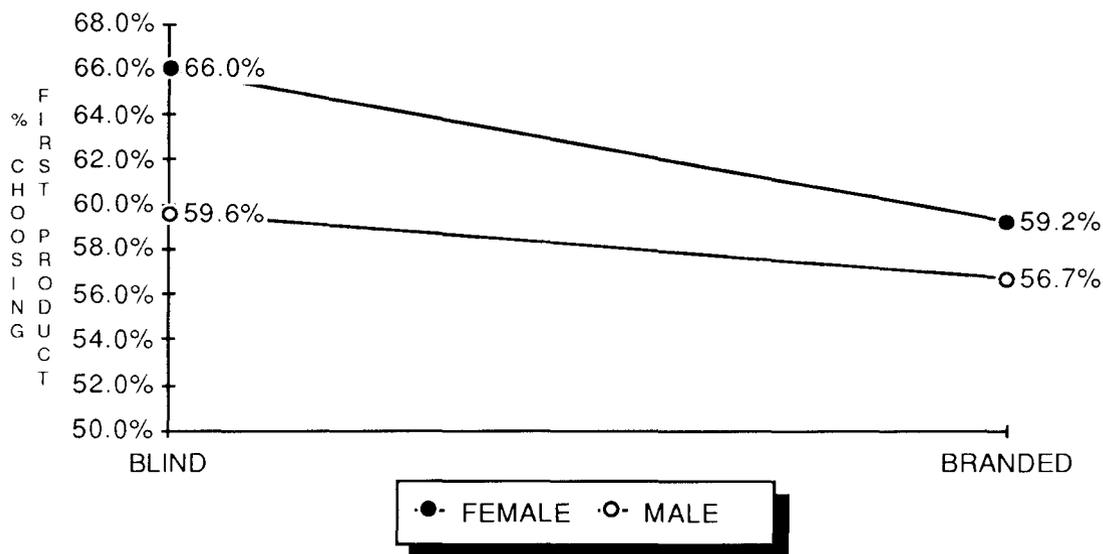
- |   |                                  |    |                                  |
|---|----------------------------------|----|----------------------------------|
| 1 | More refreshing                  | 10 | Like better overall              |
| 2 | Stronger flavor                  | 11 | Rather buy                       |
| 3 | Better flavor                    | 12 | Better aroma                     |
| 4 | Sweeter                          | 13 | Sticks more to teeth/dental work |
| 5 | Moistens mouth better            | 14 | Made mouth feel fresher          |
| 6 | More refreshing flavor           | 15 | Softer                           |
| 7 | Longer lasting taste             | 16 | Better texture                   |
| 8 | Cooled mouth more                |    |                                  |
| 9 | Better for freshening the breath |    |                                  |

Table 8

P-values for Interactions among Branding, Sex, and Age

	<u>Branding</u> <u>x Sex</u>	<u>Branding</u> <u>x Age</u>	<u>Sex</u> <u>x Age</u>	<u>Branding</u> <u>x Sex x Age</u>
Like better overall (H)	.040	.359	>.500	.424
Better flavor (H)	.194	.445	>.500	.171
Better texture (H)	.057	>.500	>.500	.215
Longer lasting taste (S)	>.500	>.500	>.500	.450
More refreshing	>.500	.237	>.500	>.500
Softer (S)	>.500	.208	>.500	>.500
Sweeter (S)	.179	.221	>.500	.279
Better aroma (H)	.236	>.500	.272	.067
More refreshing flavor	.435	>.500	>.500	>.500
Stronger flavor (S)	.280	.428	>.500	>.500
Cooled mouth more (S)	.277	.457	>.500	.435
Made mouth fresher (S)	.305	>.500	>.500	>.500
Sticks more to teeth (S)	.466	.465	>.500	.424
Moistens mouth better	.159	>.500	.199	>.500
Better for fresh breath	.228	.322	>.500	.464
Rather buy (H)	.099	.330	.423	.229

Figure 2

Branding by Sex Interaction on Like Better Overall

of about ten points (that is, ten points above the 50% expected by chance). This group includes the five judgements identified as hedonic and three of the seven sensory judgements: Cooled mouth more, Made mouth feel fresher, and Longer lasting taste. The second tier consists of three sensory judgements (Stronger flavor, Softer, Sweeter) and Moistens mouth better and is associated with an order effect of four to six points. The final tier is the Sticks to teeth more judgement which is the only judgement exhibiting a second product bias (also about five points).

To test for the significance of differences between hedonic and sensory judgements, a series of two dependent variable multinomial MANOVAs pairing each hedonic judgement with each sensory judgement was run. The resulting

p-values are shown in Table 10. Because the order bias favored the second product on the Sticks more judgement, the proportion choosing the second product on Sticks more was compared with the proportion choosing the first product on the hedonic judgements. The data show that the sensory judgements in the second and third tier of Table 9 (Stronger flavor, Softer, Sweeter, and Sticks more) each exhibited significantly less order bias than each of the hedonic judgements. The remaining sensory judgements (Cooled mouth more, Made mouth feel fresher, and Longer lasting taste), however, while exhibiting significantly less bias than Better aroma, were not found to be less prone to order bias than the majority of the hedonic judgements.

#### Hypothesis 5

The fifth hypothesis was that the effect of product branding, respondent sex, and respondent age will be stronger on hedonic judgements than on sensory judgements. Tests of the effect of these factors and interactions among them on the difference between sensory and hedonic judgement order bias were conducted for each of the 35 sensory-hedonic variable pairs described above. In the section presented below, the hypothesis stated in the null form (There is no interaction between type of judgement and product branding, type of judgement and sex, and type of judgement and age.) is addressed for each factor separately.

#### Branding

Significant differences ( $p < .05$ ) between sensory and hedonic judgements in size of effects related to product branding are summarized in

Table 9

Ranking of 16 Paired Comparison Judgements on Proportion Choosing First Product

	<u>Total Sample</u> (n=8000)	<u>Order bias</u> <u>P-value</u>
1. Better aroma (H)	64.1%	<.001
2. Rather buy (H)	61.3	<.001
3. More refreshing	61.2	<.001
4. Like better overall (H)	60.8	<.001
5. More refreshing flavor	60.7	<.001
6. Cooled mouth more (S)	60.7	<.001
7. Better for freshening the breath	60.6	<.001
8. Made mouth feel fresher (S)	60.5	<.001
9. Longer lasting taste (S)	60.4	<.001
10. Better flavor (H)	60.1	<.001
11. Better texture (H)	59.8	<.001
12. Moistens mouth better	56.2	<.001
13. Stronger flavor (S)	55.8	<.001
14. Softer (S)	54.0	<.001
15. Sweeter (S)	53.8	<.001
16. Sticks more to teeth or dental work (S)	44.5	<.001

Note. n slightly smaller on judgements other than "Like better overall" due to small number of "no preference" responses.

Table 10

P-values associated with Differences in Proportion Choosing First Product  
between Sensory and Hedonic Judgements for 35 Sensory-Hedonic  
Judgement Pairs

	<b>Hedonic judgements</b>				
	<u>Better Aroma</u>	<u>Rather Buy</u>	<u>Like better Overall</u>	<u>Better Flavor</u>	<u>Better Texture</u>
<b><u>Sensory judgements</u></b>					
Cooled mouth more	p<.001 *	p=.195	p>.500	p=.237	p=.175
Made mouth feel fresher	p<.001 *	p=.052	p=.385	p=.279	p=.169
Longer lasting taste	p<.001 *	p=.043 *	p=.320	p=.470	p=.362
Stronger flavor	p<.001 *	p<.001 *	p<.001 *	p<.001 *	p<.001 *
Softer	p<.001 *	p<.001 *	p<.001 *	p<.001 *	>p<.001 *
Sweeter	p<.001 *	p<.001 *	p<.001 *	p<.001 *	p<.001 *
Sticks more to teeth /dental work	p<.001 *	p=.003 *	p<.001 *	p<.001 *	p<.001 *
* p<.05					

Note. Responses for Sticks more judgement reversed for significance test due to second product bias.

Table 11. Eighteen of the 35 sensory-hedonic pairs evidenced differences in branding effect or on an interaction involving branding.

Two of the sensory judgements behaved very similarly to the hedonic judgements with respect to product branding. The effect of branding on Cooled mouth more order bias was not different than that on any of the five hedonic judgements. Made mouth feel fresher was generally not differently influenced by branding than the hedonic judgements. The main effect due to branding on Rather buy (64.0% blind vs. 58.5% branded) was somewhat stronger ( $p=.029$ ), however, than on Made mouth feel fresher (62.8% blind vs. 58.7% branded).

The main effect due to branding was weaker for Longer lasting taste (61.8% blind vs. 61.0% branded) than for Rather buy ( $p=.003$ ) and for Like better overall (63.4% blind vs. 58.2% branded;  $p=.009$ ). Longer lasting taste did not differ, however, from Better flavor, Better texture, or Better aroma on the effect of branding. The significant branding-sex-age interaction between Longer lasting taste and Better aroma was driven by a stronger effect of sex on Better Aroma among 18-34 year olds in the blind condition. This interaction is described fully in the results relating to sex-judgement type interactions. The interaction of branding and sex was weaker for Stronger flavor than for Like better overall, Better flavor, and Rather buy. This difference was driven by a stronger branding effect among females on these three hedonic judgements than on Stronger flavor. The effect of branding among females was stronger for Like better overall (66.0% blind vs. 59.2% branded;  $p=.002$ ), Better flavor (64.5% blind vs. 59.3% branded;  $p=.021$ ), and Rather buy (64.0% blind vs. 59.9% branded;  $p=.003$ ) than for Stronger flavor (56.4% blind vs. 55.7% branded).

Table 11

Summary of Effects involving Product Branding that Vary\* across Sensory-Hedonic Judgement Pairs

	<u>Hedonic judgements</u>				
	<u>Better Aroma</u>	<u>Rather Buy</u>	<u>Like better Overall</u>	<u>Better Flavor</u>	<u>Better Texture</u>
<u>Sensory judgements</u>					
Cooled mouth more	--	--	--	--	--
Made mouth feel fresher	--	B	--	--	--
Longer lasting taste	BSA	B	B	--	--
Stronger flavor	B,BS,BSA	B,BS	B,BS	B,BS	B,BS,BSA
Softer	--	B	B	B	--
Sweeter	B	B,BA	B,BA	B,BA	B
Sticks more to teeth/dental work	--	--	--	--	--
B = main effect due to branding					
BS = branding-sex interaction					
BA = branding-age interaction					
BSA = branding-sex-age interaction					
* p<.05					

Note. Responses for Sticks more judgement reversed for significance test due to second product bias.

The branding-sex-age interaction on the difference between Stronger flavor and Better texture showed a similar pattern, with the qualification that the branding effect was stronger for the hedonic judgement among younger women only (12-17 and 18-34). The effect of branding on Better texture among 12-17 year old females (63.6% blind vs. 54.9% branded, with greater order bias blind) was significantly different ( $p=.002$ ) than on Stronger flavor (49.1% blind vs. 54.4% branded, with greater order bias branded). Among 18-34 year old females, while both judgements showed greater first product bias in the blind condition, the effect of branding was larger ( $p=.050$ ) for Better texture (64.0% blind vs. 58.6% branded) than for Stronger flavor (57.2% blind vs. 56.8% branded).

The effect of branding on Stronger flavor was weaker than on Better aroma in some groups. Females 12-17 and males 35-55 showed greater order bias in the blind condition on Better aroma (66.9% and 64.7% blind vs. 57.2% and 58.2% branded). The effect of branding on Stronger flavor was significantly different than on Better Aroma for females 12-17 ( $p=.002$ ), with greater first product bias branded than blind (49.1% blind vs. 54.4% branded) and for males 35-55 ( $p=.017$ ) with negligible branding effect (56.2% blind vs. 57.2% branded).

Order bias on Softer was relatively unaffected by branding (54.3% blind vs. 53.7% branded). The effect of branding was significantly greater on three of the five hedonic judgements, all with greater bias blind: Like better overall (63.4% blind vs. 58.2% branded;  $p=.014$ ), Better flavor (62.4% blind vs. 57.8% branded;  $p=.034$ ), Rather buy (64.0% blind vs. 58.5% branded;  $p=.006$ ).

As with Softer, branding displayed no significant effect on Sweeter (54.6% blind vs. 53.0% branded). As predicted, main effects due to branding

were stronger on Better aroma (66.0% blind vs. 62.3% branded;  $p=.013$ ) and Better texture (61.8% blind vs. 57.7% branded;  $p=.031$ ). Differences between Sweeter and the other three hedonic judgements in effect of branding were limited primarily to younger respondents. Sweeter showed no significant branding effect among 12-17 year olds (48.1% blind vs. 50.5% branded) and among 18-34 year olds (54.7% blind vs. 52.9% branded). Among 12-17 year olds, Like better overall (62.7% blind vs. 53.3% branded;  $p<.001$ ) and Better flavor (60.4% blind vs. 51.9% branded;  $p=.002$ ) showed main effects due to branding larger than Sweeter. Rather buy showed significantly greater branding effect than Sweeter among both 12-17 year olds (63.3% blind vs. 53.7% branded;  $p=.001$ ) and 18-34 year olds (64.1% blind vs. 58.7% branded;  $p=.026$ ).

As noted earlier, Sticks to teeth/dental work more showed a second product bias. Choice of the first product was lower in the blind condition (3.9%). The effect of branding on order bias for Sticks more was not significantly different than for the five hedonic judgements.

### Sex

Significant differences ( $p<.05$ ) between sensory and hedonic judgements in effects related to the sex of the respondent are summarized in Table 12. Seventeen of the 35 sensory-hedonic pairs showed differences in sex effect or on interactions involving sex.

As was found for branding, two of the sensory judgements behaved very similarly to the hedonic judgements with respect to respondent sex. The effect of sex on Cooled mouth more order bias was not different than its effect on four

Table 12

Summary of Effects involving Respondent Sex that Vary\* across Sensory-Hedonic Judgement Pairs

	<b>Hedonic judgements</b>				
	<u>Better Aroma</u>	<u>Rather Buy</u>	<u>Like better Overall</u>	<u>Better Flavor</u>	<u>Better Texture</u>
<b><u>Sensory judgements</u></b>					
Cooled mouth more	--	--	--	--	S
Made mouth feel fresher	--	--	--	--	--
Longer lasting taste	BSA	SA	SA	--	--
Stronger flavor	BS,BSA	BS	BS	BS	BS,BSA
Softer	S	S	S	S	--
Sweeter	S	S	S	S	--
Sticks more to teeth/dental work	--	--	--	--	--
S = main effect due to branding					
BS = branding-sex interaction					
SA = sex-age interaction					
BSA = branding-sex-age interaction					
* p<.05					

Note. Responses for Sticks more judgement reversed for significance test due to second product bias.

of the five hedonic judgements, with all judgements showing greater order bias among females. The sex effect for Cooled mouth more (62.9% female vs. 57.4% male) was actually significantly stronger ( $p=.039$ ) than that for Better texture (61.0% female vs. 58.0% male). This difference was in the opposite direction of that predicted. The effect of sex on Made mouth feel fresher was not significantly different than on any of the hedonic judgements.

Order bias on Longer lasting taste showed a smaller sex effect than some hedonic judgements as was hypothesized, but only among 18-34 year olds. The main effect due to sex was weaker among 18-34 year olds for Longer lasting taste (61.8% female vs. 59.4% male) than for Rather buy (63.9% female vs. 57.7% male;  $p=.003$ ) and for Like better overall (63.2% female vs. 57.5% male;  $p=.003$ ). A similar phenomenon was observed for Better aroma among 18-34 year olds, but only in the blind condition (Longer lasting taste: 62.4% female/61.3% male vs. Better aroma: 69.8% female/61.8% male;  $p=.003$ ). Longer lasting taste did not differ from Better flavor or Better texture on the effect of sex.

The effect of respondent sex on order bias on Stronger flavor was significantly weaker in the blind condition (56.4% female vs. 55.1% male) than on Better aroma (68.4% female vs. 62.6% male;  $p=.030$ ), Rather buy (64.0% female vs. 60.5% male;  $p=.008$ ), Like better overall (66.0% female vs. 59.6% male;  $p=.008$ ), and Better flavor (64.5% female vs. 59.5% male;  $p=.021$ ). However, in the branded condition among 12-17 year olds, the effect of sex was actually greater for Stronger flavor (54.4% female vs. 45.7% male) than for Better aroma (57.2% female vs. 61.1% male;  $p=.016$ ) and Better texture (54.9% female vs. 59.1% male;  $p=.010$ ).

Order bias on Softer was unaffected by respondent sex (first product selection 54% among both males and females). The effect of respondent sex was significantly greater on four of the five hedonic judgements, all with greater bias among females: Better aroma (66.3% female vs. 61.0% male;  $p=.040$ ), Rather buy (63.2% female vs. 58.5% male;  $p=.048$ ), Like better overall (62.6% female vs. 58.1% male;  $p=.047$ ), and Better flavor (61.9% female vs. 57.5% male;  $p=.024$ ). Better texture did not show significantly greater sex effect (61.0% female vs. 58.0% male;  $p=.190$ ).

As with Softer, sex displayed no significant effect on Sweeter order bias (54.0% female vs. 53.5% male). Better aroma (66.3% female vs. 61.0% male;  $p=.032$ ), Rather buy (63.2% female vs. 58.5% male;  $p=.028$ ), Like better overall (62.6% female vs. 58.1% male;  $p=.047$ ), and Better flavor (61.9% female vs. 57.5% male;  $p=.022$ ) all exhibited stronger sex effects.

As noted earlier, Sticks to teeth/dental work more showed a second product bias. The effect of sex on order bias was not different for this judgement than on any of the hedonic judgements.

### Age

Significant differences between sensory and hedonic judgements in effects related to the age of the respondent are summarized in Table 13. Thirteen of the 35 sensory-hedonic pairs showed differences in age effects or on interactions involving age.

The effect of age on order bias for Cooled mouth more and Made mouth feel fresher was not different than its effect on the five hedonic judgements. This is not surprising, as neither of these judgements independently show significant

Table 13

Summary of Effects involving Respondent Age that Vary\* across Sensory-Hedonic Judgement Pairs

	<u>Hedonic judgements</u>				
	<u>Better Aroma</u>	<u>Rather Buy</u>	<u>Like better Overall</u>	<u>Better Flavor</u>	<u>Better Texture</u>
<b><u>Sensory judgements</u></b>					
Cooled mouth more	--	--	--	--	--
Made mouth feel fresher	--	--	--	--	--
Longer lasting taste	BSA	SA	SA	--	--
Stronger flavor	A,BSA	A	A	--	A,BSA
Softer	--	--	--	A	--
Sweeter	--	BA	BA	BA	--
Sticks more to teeth/dental work	A	--	--	A	--

S = main effect due to branding

BS = branding-sex interaction

SA = sex-age interaction

BSA = branding-sex-age interaction

\*  $p < .05$

Note. Responses for Sticks more judgement reversed for significance test due to second product bias.

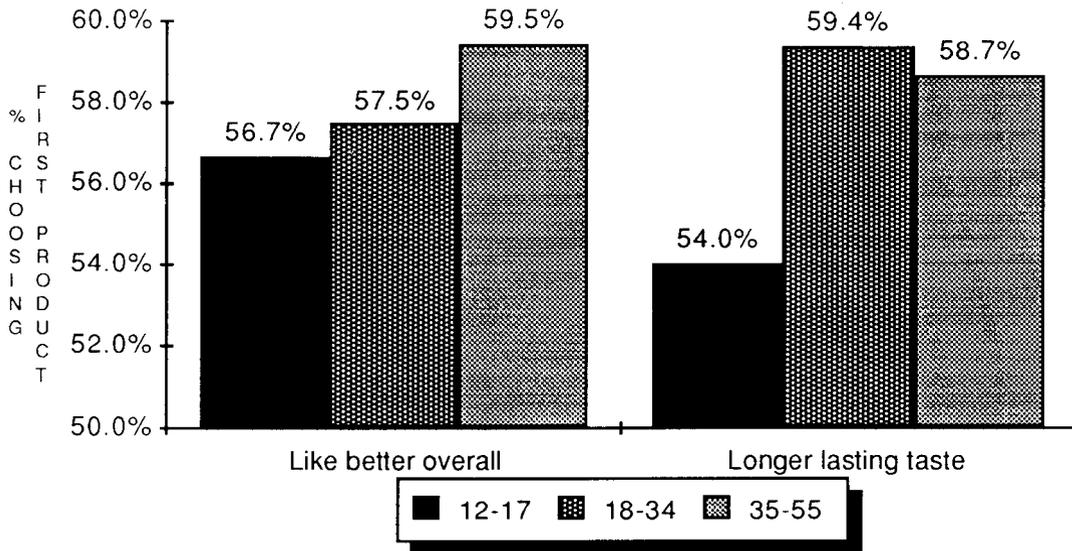
age effects.

The effect of age on Longer lasting taste order bias was also generally not different than that on the hedonic judgements. An exception is the larger effect age has on Longer lasting taste than on Like better overall among males ( $p=.034$ ). As is shown in Figure 3, the difference is driven by the greater attenuation of order bias associated with membership in the 12-17 age group for Longer lasting taste. This finding is in the opposite direction to that hypothesized. Among females in the blind condition (Figure 4) the age effect is different on Better aroma than Longer lasting taste ( $p=.046$ ). This difference, however, cannot be characterized as "stronger" for either judgement.

Stronger flavor was one of the judgements on which order bias varied significantly by age ( $p=.001$ ), with no order bias among 12-17 year olds (49.6% first product choice) versus a six to seven point first product bias among 18-34 year olds (56.2%) and 35-55 year olds (57.7%). Comparison with age effects on the hedonic judgements shows that the hypothesis is not supported for Stronger flavor. With the exception of Better flavor ( $p=.204$ ), the hedonic judgements all showed significantly less rather age effect than Stronger flavor: Better aroma ( $p=.016$ ), Rather buy ( $p=.018$ ), Like better overall ( $p=.026$ ), and Better flavor ( $p=.004$ ). Inspection of the two branding-sex-age interactions reveals that the contrast is even greater among males in the branded condition with Stronger flavor showing greater response to age than Better aroma ( $p=.005$ ) and Better texture ( $p=.004$ ).

Order bias on Softer was unaffected by respondent age (first product selection 53.7% for 12-17 year olds, 54.9% for 18-34 year olds, 53.0% for 35-55 year olds). However, only one of the hedonic judgements displayed a

Figure 3

Age Effect among Males for Like Better Overall vs. Longer Lasting Taste

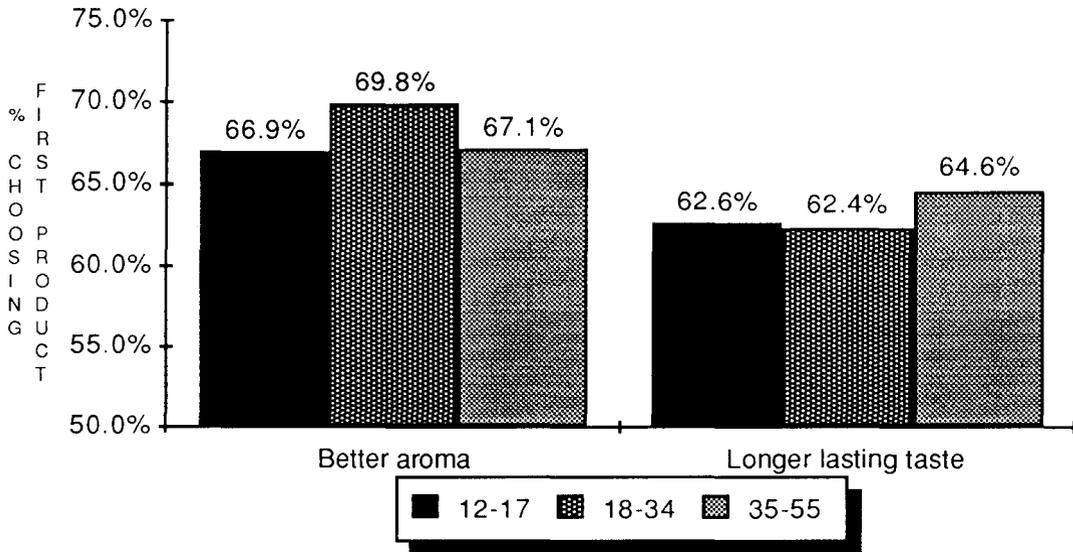
significantly greater relationship between age and order bias. The age effect on Better flavor (first product selection 56.1% for 12-17 year olds, 59.8% for 18-34 year olds, 62.0% for 35-55 year olds) was significantly stronger ( $p=.004$ ) than on Softer with 12-17 year olds showing less order bias than older respondents.

Sweeter was one of the judgements for which order bias varied by age (first product selection 49.3% for 12-17 year olds, 53.8% for 18-34 year olds, 59.5% for 35-55 year olds). Compared with the hedonic judgements, this effect was stronger than Like better overall ( $p=.016$ ) and Rather buy ( $p=.012$ ). This difference is in the opposite direction of that hypothesized.

The effect of age on Sticks more was significantly different than on Better aroma and Better flavor. As Figure 5 shows, order bias for Sticks more was

Figure 4

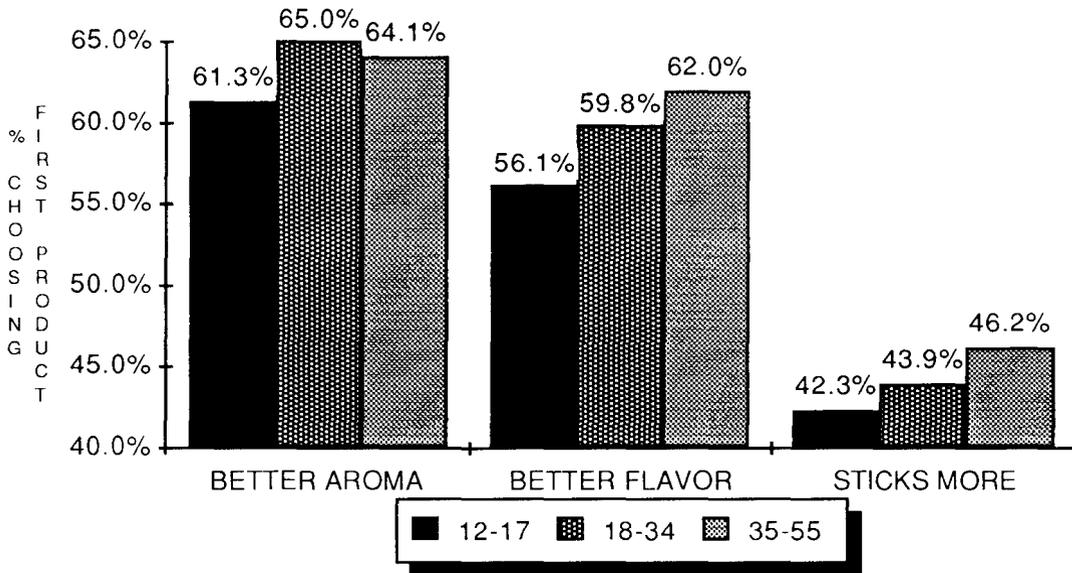
Age Effect among Females Tasting Blind Products for Better Aroma vs. Longer Lasting Taste



greater among younger respondents while order bias for Better flavor was greater among older respondents. This age-judgement type interaction is disordinal, with magnitude of the effects similar but in opposite directions.

Figure 5

Percent Choosing First Product by Age for Sticks More vs. Better Aroma and Better Flavor



## CHAPTER 5

### DISCUSSION

#### Hypothesis 1

The first hypothesis (Order effects are stronger for unbranded product pairs than for branded product pairs.) was supported for Like better overall and the majority of the other judgements. The effect size ranged from a little over five percentage points for Like better overall to no effect on several of the sensory judgements.

There are two potential phenomena that might explain why respondents in the branded condition showed less order bias. First, the interpretation of sensory information may be easier when a context in the form of a brand name associated with prior experience is provided. This reasoning is consistent with the work of Day (1969) and Mitchell (1956) that suggested that order effects are associated with respondent difficulty in discriminating between products. Second, respondents faced with two similar products may use non-sensory information external to the test such as attitudes toward the brands to make a decision.

If lower order bias can be primarily attributed to the first explanation, an argument could be made for the advantage of branded over blind testing in reducing error variance. If, however, the lowered order bias in the branded condition is primarily due to the second explanation, the internal validity of

of branded testing could be questioned. Further research on the sensitivity of branded testing to physical product-based differences would shed further light on this issue.

### Hypothesis 2

The second hypothesis (Order effects are stronger for male respondents than for female respondents.) was not supported. Females showed stronger order bias than males on Like better overall and several of the other judgements. The effect sizes ranged from over five percentage points on Cooled mouth more to no significant difference on Sweeter, Softer, and Sticks to teeth/dental work more.

The reason why females showed more order bias than males is not clear. Basic research on taste perception (Wysocki and Gilbert, 1989; Cowart, 1989) shows that females have superior ability to males in some circumstances. If this is so, and Day and Mitchell's hypothesis about discrimination difficulty contributing to order bias is valid, then males would be expected to show more order bias than females.

Given that order effects were stronger for females in this study, what might account for this finding? Logically, the greater order effects found among females in the study reported here are due either to a physiological difference or to a difference in response style between males and females. As noted earlier, the physiological difference is most likely in the direction of greater sensitivity among females. It is possible that response to the first product among females is stronger than among males, rendering females less able to respond objectively to the second product. Unfortunately, evaluation of

physiologically-based theories are beyond the scope of this research project.

A second possible reason why females exhibited greater order bias than males is that females may exhibit a response style that predisposes them toward choosing the first product. A possible explanation for such a style is that the first product is perceived to be the one favored by the experimenter (by virtue of its being presented first) and females are more susceptible to this demand characteristic. If this were true, we would expect to see the effect hold up across a variety of products in further research.

### Hypothesis 3

The third hypothesis (Order effects are stronger for older respondents than for younger respondents.) was not supported for Like better overall, but was supported for teens versus adults on several of the other judgements including More refreshing, Stronger flavor, Better flavor, Sweeter, Moistens mouth better, More refreshing flavor, and Longer lasting taste. Teens exhibited four to seven points less order bias on average than adults on these judgements.

Why would teens show lower order bias? Again, the findings could stem from a sensory/physiological basis or from a response style. Given the age range represented in this study and basic research findings supporting sensory response differences due to age, it is unlikely that this effect is due to teens sensing the stimuli differently than adults. It is more likely that teens have a different response style than adults. One explanation consistent with the results is that teens are less likely to be committed to the first product and thus may exhibit less order bias. It is generally known that teens like to experiment and

try a number of different products (Rand, 1988). Chewing gum usage studies show that the average teenager has chewed a greater number of different brands and flavors of gum in the past week than the average adult.

However, response style does not explain why 12-17 year olds would show less order bias than adults on some judgements but not on others. Sweeter and Stronger flavor are two of the judgements where 12-17 year olds showed less order bias than adults. Children are known to have a greater interest in sweeter, stronger flavored foods than adults. It may be that teens have better discrimination ability with respect to sweetness and flavor strength because these attributes are more salient for them.

#### Hypothesis 4

The fourth hypothesis (Order effects are stronger with hedonic-type judgements than with sensory-type judgements.) was not supported overall. Three of the seven sensory-type judgements showed order bias that was not significantly different than the order bias found for the majority of the hedonic judgements. These three judgements are: Cooled mouth more, Made mouth feel fresher, and Longer lasting taste. The other four sensory judgements showed significantly less order bias than the five hedonic judgements. These judgements were: Stronger flavor, Softer, Sweeter, and Sticks more to teeth.

Examination of the subset of the sensory judgements that did not behave in the hypothesized manner vis a vis the hedonic judgements reveals that they all have a unipolar evaluative dimension. That is, cooling, mouth freshening, and long lasting taste are all characteristics for which more is usually better in a chewing gum. In fact, examination of the correlation between each of these

judgements and Like better overall shows that they are all strongly positively associated (.7 and higher) with preference (Table 14).

Table 14

Correlations (Phi) between Sensory Judgements and Like Better Overall

---

	<u>Like better overall</u>
<b><u>Sensory judgements</u></b>	
Cooled mouth more	.63
Made mouth feel fresher	.74
Longer lasting taste	.72
<hr/>	
Stronger flavor	.50
Softer	.27
Sweeter	.35
Sticks more to teeth	
/dental work	.23

---

The four judgements that did show lower order bias compared to the hedonic judgements had considerably lower correlations with Like better overall (i.e., .5 and lower). On these measures, unlike the other sensory measures, judgement suggests that more is not always better. That is, it is

possible for a gum to be too strong, too sweet, too soft, and certainly to stick to one's teeth too much.

Based on these findings, it appears that the judgement typology that best differentiates judgements into high order bias and low order bias is "evaluative" versus "non-evaluative". Further research with other evaluative and non-evaluative judgements would help confirm this hypothesis.

### Hypothesis 5

The fifth hypothesis (Product branding and respondent age and sex interact with judgement-type such that their effect on order bias is strongest for hedonic-type judgements and weakest for sensory-type judgements.) was supported for product branding and sex on a subset of the sensory judgements but not for respondent age. Whether the products were sampled blind or branded generally made more of a difference on the hedonic judgements than on three of the four non-evaluative sensory judgements identified above (Stronger flavor, Softer, and Sweeter). The effect of branding on Sticks more, Cooled mouth more, Made mouth feel fresher, and Longer lasting taste was generally not different than on the hedonic judgements.

As with hypothesis 4, the failure to find a lesser branding effect on Cooled mouth more, Made mouth feel fresher, and Longer lasting taste than on the hedonic judgements may be due to the evaluative nature of these judgements. It should be noted that for the Sticks more judgement the power of the statistical test was weakened by non-response by close to half of respondents (46.9%).

As hypothesized, sex effects were generally weaker on Stronger flavor, Softer, and Sweeter than for the hedonic judgements. Again, the other sensory judgements did not behave as hypothesized, possibly stemming from their evaluative nature.

The age variable compared to the branding and sex variables were found to interact differently with the judgements. Most of the sensory judgements did not differ from the hedonic judgements in the effect of age on order bias. Where differences existed, age tended to have more of an effect on the sensory judgements, a difference opposite to that predicted. Stronger flavor and Sweeter showed a stronger relationship between age and order bias than several of the hedonic judgements. As reported in the section describing the findings related to hypothesis 3, where age and judgement-type interactions exist they tend to be driven by the 12-17 year old age group showing less order bias than adults.

In summary, the effect of product branding and respondent sex on order bias tends to be greater on hedonic judgements than on a subset of the sensory judgements. This subset comprises those sensory judgements that do not have a clear unipolar evaluative component associated with their use.

### Implications for practice

Overall, the results of this study show that order bias in paired comparison taste tests is strongly influenced by whether the product is presented blind or branded, the age and sex of the respondent, and the type of judgements used. In terms of practice, branded testing appears to have a definite advantage over blind testing in that it significantly reduces order bias.

In cases where it can be established that there are no significant threats to the internal validity of a branded test, it is recommended that it be used over blind testing.

Female respondents are more susceptible to order bias than males. When testing among females, larger samples may be necessary than when testing among males to overcome this source of order bias. Teen response to paired comparison tasks as contrasted with adults needs further study.

Finally, in cases where non-evaluative judgements can be used to meet test objectives they are preferable to hedonic and other evaluative judgements due to lower order bias.

#### Limitations

A number of limitations of the study must be considered with respect to utilizing the information presented. One major limitation of this study is that it was conducted on a single product category, chewing gum. An attempt should be made to determine whether the findings can be generalized to other product categories. A second limitation is that the different judgements were measured using the same subjects. That is to say, the measures were not independent. It is not clear to what extent order bias would be present for each judgement if it were presented as a single task independent of the other judgements. It is likely that in the study presented here respondents' judgements on some judgements were influenced by their response to the Like better overall judgement that was presented first. A final limitation of this study is that it is limited to describing only the correlates of order bias and does not directly address the causal mechanisms underlying the phenomenon. Nevertheless,

an understanding of the correlates of presentation order bias can be useful when developing hypotheses related to causality.

APPENDIX 1  
QUESTIONNAIRES

Peryam & Kroll Research Corporation  
Chicago, Illinois

GUM QUESTIONNAIRE

1. In the past seven days, how many pieces of chewing gum have you chewed? Do not include bubble gum.

NONE                      1-4                      5-10                      11 OR MORE  
(IF "NONE," TERMINATE).

2. I will mention several different flavors of gum. Please tell me if you like or dislike the following flavors.

Cinnamon	LIKE	DISLIKE	Peppermint	LIKE	DISLIKE
Spearmint	LIKE	DISLIKE	Wintergreen	LIKE	DISLIKE

(IF "DISLIKE" TO "SPEARMINT," TERMINATE)

3. I will mention some age ranges. Please tell me where your age falls.

UNDER 12              12-17              18-34              35-55              54 AND OVER  
(IF "UNDER 12" OR "56 AND OVER," TERMINATE).  
(CHECK QUOTA)

4. What brands and flavors of chewing gum have you chewed in the past seven days?  
(HAND DISPLAY TO RESPONDENT, ASK RESPONDENT TO GIVE NUMBER). USE  
DISPLAY FOR Q.4 & Q.5.

5. What is the one brand and flavor of gum you chew most often? (USE SAME DISPLAY FOR  
Q.4, ASK RESPONDENT TO GIVE NUMBER).

\_\_\_\_\_ (SEE QUALIFICATION PAGE)

6. Are you or anyone in your family employed in any of the following industries?

Advertising Agency	YES	NO
Marketing Research	YES	NO
Food, Chewing Gum or Confectionery Manufacturer	YES	NO

(IF "YES" TO ANY, TERMINATE)

7. How often do you have trouble with gum sticking to your teeth?

ALWAYS      FREQUENTLY      OCCASIONALLY      SELDOM      NEVER

8. You are invited to participate in a taste test where you will chew two samples of gum. One or both of the samples may contain saccharin or aspartame. Would you have any objections to this?

YES                      NO  
(IF "YES," TERMINATE)

9. Think about how long you normally chew a stick of gum. I will mention various lengths of time in minutes. Please tell me which is closest to your normal chewing time. (READ IN ASCENDING ORDER AND CIRCLE ANSWER).

5      10      15      20      25      30      35      40      45      55      over 55

- 1 2 (CHECK QUOTA)

GUM RATING FORM

Serial # \_\_\_\_\_

BEFORE YOU START CHEWING, PLEASE ANSWER THESE QUESTIONS.

<u>APPEARANCE</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>AROMA (SMELL)</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>SIZE</u>	TOO SMALL		ABOUT RIGHT		TOO LARGE
<u>COLOR</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR

NOW START CHEWING THE GUM AND ANSWER THESE QUESTIONS WHEN YOU FIRST BITE INTO IT.

<u>TEXTURE</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>TEXTURE</u>	TOO SOFT		ABOUT RIGHT		TOO HARD

AFTER ONE MINUTE ANSWER THESE QUESTIONS.

<u>FLAVOR STRENGTH</u>	TOO WEAK		ABOUT RIGHT		TOO STRONG
<u>SWEETNESS</u>	NOT SWEET ENOUGH		ABOUT RIGHT		TOO SWEET
<u>TEXTURE</u>	TOO SOFT		ABOUT RIGHT		TOO FIRM
<u>AMOUNT OF COOLING</u>	NOT ENOUGH		ABOUT RIGHT		TOO MUCH
<u>FRESHENING THE BREATH</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FLAVOR</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>TEXTURE</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FRESHENS THE MOUTH</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FLAVOR LEVEL</u>	NONE	SLIGHT	DEFINITE	STRONG	VERY STRONG

HOW MUCH DOES THIS GUM STICK TO YOUR TEETH OR DENTAL WORK?

NOT AT ALL    VERY SLIGHTLY    SLIGHTLY    MODERATELY    A GREAT DEAL

AFTER THREE MINUTES ANSWER THESE QUESTIONS.

<u>FLAVOR STRENGTH</u>	TOO WEAK		ABOUT RIGHT		TOO STRONG
<u>SWEETNESS</u>	NOT SWEET ENOUGH		ABOUT RIGHT		TOO SWEET
<u>TEXTURE</u>	TOO SOFT		ABOUT RIGHT		TOO FIRM
<u>AMOUNT OF COOLING</u>	NOT ENOUGH		ABOUT RIGHT		TOO MUCH
<u>FRESHENING THE BREATH</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FLAVOR</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>TEXTURE</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FRESHENS THE MOUTH</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FLAVOR LEVEL</u>	NONE	SLIGHT	DEFINITE	STRONG	VERY STRONG

HOW MUCH DOES THIS GUM STICK TO YOUR TEETH OR DENTAL WORK?

NOT AT ALL    VERY SLIGHTLY    SLIGHTLY    MODERATELY    A GREAT DEAL

AFTER SIX MINUTES ANSWER THESE QUESTIONS.

<u>FLAVOR STRENGTH</u>	TOO WEAK		ABOUT RIGHT		TOO STRONG
<u>SWEETNESS</u>	NOT SWEET ENOUGH		ABOUT RIGHT		TOO SWEET
<u>TEXTURE</u>	TOO SOFT		ABOUT RIGHT		TOO FIRM
<u>AMOUNT OF COOLING</u>	NOT ENOUGH		ABOUT RIGHT		TOO MUCH
<u>FRESHENING THE BREATH</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FLAVOR</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>TEXTURE</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FRESHENS THE MOUTH</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FLAVOR LEVEL</u>	NONE	SLIGHT	DEFINITE	STRONG	VERY STRONG

HOW MUCH DOES THIS GUM STICK TO YOUR TEETH OR DENTAL WORK?

NOT AT ALL    VERY SLIGHTLY    SLIGHTLY    MODERATELY    A GREAT DEAL

AFTER NINE MINUTES ANSWER THESE QUESTIONS.

<u>FLAVOR STRENGTH</u>	TOO WEAK		ABOUT RIGHT		TOO STRONG
<u>SWEETNESS</u>	NOT SWEET ENOUGH		ABOUT RIGHT		TOO SWEET
<u>TEXTURE</u>	TOO SOFT		ABOUT RIGHT		TOO FIRM
<u>AMOUNT OF COOLING</u>	NOT ENOUGH		ABOUT RIGHT		TOO MUCH
<u>FRESHENING THE BREATH</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FLAVOR</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>TEXTURE</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FRESHENS THE MOUTH</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FLAVOR LEVEL</u>	NONE	SLIGHT	DEFINITE	STRONG	VERY STRONG

HOW MUCH DOES THIS GUM STICK TO YOUR TEETH OR DENTAL WORK?

NOT AT ALL    VERY SLIGHTLY    SLIGHTLY    MODERATELY    A GREAT DEAL

AFTER TWELVE MINUTES ANSWER THESE QUESTIONS.

<u>FLAVOR STRENGTH</u>	TOO WEAK		ABOUT RIGHT		TOO STRONG
<u>SWEETNESS</u>	NOT SWEET ENOUGH		ABOUT RIGHT		TOO SWEET
<u>TEXTURE</u>	TOO SOFT		ABOUT RIGHT		TOO FIRM
<u>AMOUNT OF COOLING</u>	NOT ENOUGH		ABOUT RIGHT		TOO MUCH
<u>FRESHENING THE BREATH</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FLAVOR</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>TEXTURE</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FRESHENS THE MOUTH</u>	EXCELLENT	GOOD	AVERAGE	POOR	VERY POOR
<u>FLAVOR LEVEL</u>	NONE	SLIGHT	DEFINITE	STRONG	VERY STRONG
<u>SIZE</u>	TOO SMALL		ABOUT RIGHT		TOO LARGE

HOW MUCH DOES THIS GUM STICK TO YOUR TEETH OR DENTAL WORK?

NOT AT ALL    VERY SLIGHTLY    SLIGHTLY    MODERATELY    A GREAT DEAL

COMPARED TO THE GUM YOU CHEW MOST OFTEN, HOW LONG DID THE FLAVOR OF THIS GUM LAST?

MUCH LONGER	A LITTLE LONGER	ABOUT THE SAME	A LITTLE SHORTER	MUCH SHORTER
----------------	--------------------	-------------------	---------------------	-----------------

HOW REFRESHING WAS THIS GUM?

EXTREMELY    MODERATELY    SLIGHTLY    VERY SLIGHTLY    NOT AT ALL

EVERYTHING CONSIDERED, HOW DID YOU LIKE THIS GUM?

LIKE EXTREMELY  
LIKE VERY MUCH  
LIKE MODERATELY  
LIKE SLIGHTLY  
NEITHER LIKE NOR DISLIKE  
DISLIKE SLIGHTLY  
DISLIKE MODERATELY  
DISLIKE VERY MUCH  
DISLIKE EXTREMELY

WHAT DO YOU LIKE ABOUT THIS GUM?

WHAT DO YOU DISLIKE ABOUT THIS GUM?

IF THIS GUM WERE IN THE STORES WOULD YOU:

DEFINITELY  
BUY IT

PROBABLY  
BUY IT

MIGHT OR MIGHT  
NOT BUY IT

PROBABLY  
NOT BUY IT

DEFINITELY  
NOT BUY IT

OVERALL, HOW MUCH DID THIS GUM STICK TO YOUR TEETH OR DENTAL WORK?

NOT AT ALL

VERY SLIGHTLY

SLIGHTLY

MODERATELY

A GREAT DEAL

Peryam & Kroll Research Corporation  
Chicago, Illinois

COMPARISON QUESTIONNAIRE

SERIAL NO. \_\_\_\_\_

WRITE IN THE BRAND OF YOUR ANSWER IN THE BLANK.

WHICH GUM DID YOU LIKE BETTER, EVERYTHING CONSIDERED? \_\_\_\_\_

WHY DID YOU LIKE THIS GUM BETTER?  
\_\_\_\_\_  
\_\_\_\_\_

WHICH GUM HAD THE BETTER FLAVOR? \_\_\_\_\_

WHICH GUM HAD THE BETTER TEXTURE? \_\_\_\_\_

WHICH GUM HAD THE LONGER LASTING TASTE? \_\_\_\_\_

WHICH GUM WAS MORE REFRESHING \_\_\_\_\_

WHICH GUM WAS SOFTER? \_\_\_\_\_

WHICH GUM WAS SWEETER? \_\_\_\_\_

WHICH GUM HAD THE BETTER AROMA (SMELL)? \_\_\_\_\_

WHICH GUM HAD MORE REFRESHING FLAVOR? \_\_\_\_\_

WHICH GUM HAD THE STRONGER FLAVOR? \_\_\_\_\_

WHICH GUM COOLED YOUR MOUTH MORE? \_\_\_\_\_

WHICH GUM MADE YOUR MOUTH FEEL FRESHER? \_\_\_\_\_

WHICH GUM STICKS MORE TO YOUR TEETH OR DENTAL WORK? \_\_\_\_\_

WHICH GUM MOISTENS YOUR MOUTH BETTER? \_\_\_\_\_

WHICH GUM WAS BETTER FOR FRESHENING THE BREATH? \_\_\_\_\_

WHICH GUM WOULD YOU RATHER BUY? \_\_\_\_\_

Name \_\_\_\_\_

Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip \_\_\_\_\_

Phone (    ) \_\_\_\_\_

APPENDIX 2  
LEVEL OF "NO PREFERENCE" RESPONSE

Table 15

Level of "No preference" Response by Judgement

---

		<u>Total Sample</u> (n=8000)
1.	Like better overall (H)	0.0%
2.	Better flavor (H)	0.8%
3.	Better texture (H)	3.4%
4.	Longer lasting taste (S)	1.7%
5.	More refreshing	1.5%
6.	Softer (S)	3.4%
7.	Sweeter (S)	2.5%
8.	Better aroma (H)	2.6%
9.	More refreshing flavor	1.5%
10.	Stronger flavor (S)	1.2%
11.	Cooled mouth more (S)	1.8%
12.	Made mouth feel fresher (S)	2.3%
13.	Sticks more to teeth/dental work (S)	46.9%
14.	Moistens mouth better	5.3%
15.	Better for freshening breath	2.7%
16.	Rather buy (H)	1.4%

---

## REFERENCES

- Allan, L.G. (1984). Contingent aftereffects in duration judgments. Annals of the New York Academy of Sciences, 423, 116-130.
- Arndt, J. (1970). A beer testing experiment. The European Marketing Research Review, 5, 10-22.
- Batsell, R., & Wind, Y. (1979). Product testing: Current methods and needed developments. Journal of the Market Research Society, 22(2), 115-139.
- Berdy, D. (1979). Order effects in taste tests. Journal of the Market Research Society, 11(4), 361-371.
- Bradley, R.A. (1976). Science, statistics, and paired comparisons. Biometrics, 32, 213-232.
- Buchanan, B.S. (1987). A model for repeat trial product tests. Psychometrika, 52(1), 61-78.
- Buchanan, B.S. (1988). A model for repeat paired comparison preference tests. Psychometrika, 53(2), 209-221.
- Buchanan, B.S., Givon, M., & Goldman, A. (1987). Measurement of discrimination ability in taste tests: An empirical investigation. Journal of Marketing Research, 24, 154-163.
- Buchanan, B.S., & Smithies, R.H. (1991). Taste claims and their substantiation. Journal of Advertising Research, 31,(3), 19-35.

- Cowart, B.J. (1989). Relationships between taste and smell across the adult life span. In C. Murphy, W. S. Cain, & D.M. Hegsted (Eds.), Nutrition and the chemical senses in aging: Recent advances and current research needs. (pp. 39-55). New York: New York Academy of Sciences.
- Day, R.L. (1969). Position bias in paired product tests. Journal of Marketing Research, 6(1), 98-100.
- David, H.A. (1963). The method of paired comparisons. New York: Harper Publishing Company.
- Davidson, R.R. & Farquhar, P.H. (1976). A bibliography on the method of paired comparisons. Biometrics, 32, 241-252.
- Dean, M.L. (1980). Presentation order effects in product taste tests. Journal of Psychology, 105, 107-110.
- Dreben, E.K., Fiske, S.T., & Hastie, R. (1979). The independence of evaluative and item information: Impression and recall order effects in behavior-based impression formation. Journal of Personality and Social Psychology, 37(10), 1758-1768.
- Edwards, A.L. (1959). Manual for the Edwards Personal Preference Schedule (rev. ed.) New York: Psychological Corporation.
- Eindhoven, J., Peryam, D., Heiligman, F., & Hamman, J.W. (1964). Effects of sample sequence on food preferences. Journal of Food Science, 29(4), 520-524.
- Engen, T. (1974). Method and theory in the study of odor preference. In A Turk, J.W. Johnston, Jr., D.G. Moulton (Eds.), Human responses to environmental odors. (pp. 122-142). New York: Academic Press.
- Fagan, J.F. (1981) Infant intelligence. Intelligence, 5(3), 239-243.

- Fechner, G.T. (1860) Elemente der Psychophysik. Leipzig, Germany: Breitkopf and Hartel.
- Greenberg, A. (1958). Paired comparisons in consumer-product tests. Journal of Marketing, 22, 411-414.
- Greenberg, A. (1963). Paired comparisons vs. monadic tests. Journal of Advertising Research, 5, 44-47.
- Jamieson, D.G. & Petrusic, W.M. (1976). On a bias induced by the provision of feedback in psychophysical experiments. Acta Psychologica, 40 (3), 199-206.
- Kamen, J.M., Peryam, D.R., Peryam, D.B., & Kroll, B.J. (1969). Hedonic differences as a function of number of samples evaluated. Journal of Food Science, 34, 475-479.
- Lana, R.E. (1961). Familiarity and the order of presentation of persuasive communications. Journal of Abnormal Social Psychology, 62, 573-577.
- Lawshe, C.H., Kephart, N.C., & McCormick, E.J. (1949). The paired comparison technique for rating performance of industrial employees. Journal of Applied Psychology, 33, 69-77.
- Luce, R.D. (1959). Individual choice behavior. New York: Wiley.
- Luce, R.D. (1977). The choice axiom after twenty years. Journal of Mathematical Psychology, 15, 215-223.
- Lundgren, B. (1978). Taste discrimination versus hedonic response to sucrose in coffee beverage: An interlaboratory study. Chemical Senses and Flavor, 3(3), 249-265.

- MacKay-Soroka, S., Trehub, S.E., Bull, D.H., & Corter, C.M. (1982). Effects of encoding and retrieval conditions on infants' recognition memory. Child Development, 53(3), 815-818.
- Mitchell, J.W. (1956). Time errors in the paired comparison taste preference. Food Technology, 10, 208-210.
- Moskowitz, H.R. (1983). Product testing and sensory evaluation of foods. Westport, CT: Food & Nutrition Press, Inc.
- Moskowitz, H.R. (1985). New directions for product testing and sensory analysis of foods. Westport, CT: Food & Nutrition Press, Inc.
- Ory, J.C. (1980). Evaluative criteria: how important and to whom? CEDR Quarterly, 13, 14-16.
- Penny, J.C., Hunt, I.M., & Twyman, W.A. (1972). Product testing methodology in relation to marketing problems: A review. Journal of the Market Research Society, 14(1), 1-29.
- Rand, L. (1988, June). The baby bust and beyond: Opportunities in the youth market into the 1990s. An introduction. Paper presented at American Demographics' Outlook 1988 conference, New York.
- Schneider, B. (1988). The additivity of loudness across critical bands: A conjoint measurement approach. Perception and Psychophysics, 43(3), 211-222.
- Scowcraft, G. Consumer evaluation of flavour: Some approaches used at Campbell Soup Company. Flavour research and food acceptance. New York: Reinhold, 249-254.
- Tetlock, P.E. (1893). Accountability and the perseverance of first impressions. Social Psychology Quarterly, 46(4), 285-292.

- Thurstone, L.L. (1927). A law of comparative judgment. Psychological Review, 4, 273-286.
- van der Ven, A H.G.S. (1980). Introduction to Scaling. New York: Wiley.
- Woodward, J.A., Bonett, D.G., & Brecht, M. (1988). Introduction to Linear Models and Experimental Design. New York: Harcourt Brace Jovanovich.
- Wysocki, C.J. & Gilbert, A.N. (1989). National Geographic Smell Survey: Effects of age are heterogeneous. In C. Murphy, W. S. Cain, & D.M. Hegsted (Eds.), Nutrition and the chemical senses in aging: Recent advances and current research needs. (pp. 12-28). New York: New York Academy of Sciences.

George Hunt is Vice President, Associate Research Director at BBDO Chicago, an advertising agency in the BBDO Worldwide Network. He has been involved in work on sensory testing of products for development and claim substantiation since 1988. He received his B.A. in Engineering and Applied Science from Harvard College in 1979 and his M.Ed. in Community Counseling from Loyola of Chicago in 1984. His other work in measurement includes development of an instrument to measure perceived social support (Brown, S.D., Alpert, D., Lent, R.W., Hunt, G., & Brady, T. (1988). Perceived social support among college students: Factor structure of the Social Support Inventory. Journal of Counseling Psychology, 35(3), 472-478.). Mr. Hunt is a member of the American Marketing Association, the American Psychological Association, and the American Educational Research Association.

The dissertation submitted by George W. Hunt has been read and approved by the following committee:

Dr. Jack Kavanagh

Professor, Counseling and Educational Psychology

Loyola University of Chicago

Dr. Ronald Morgan

Professor, Counseling and Educational Psychology

Loyola University of Chicago

Dr. David Peryam

Peryam & Kroll Marketing & Sensory Research

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the dissertation is now given final approval by the Committee with reference to content and form.

The dissertation is, therefore, accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 18, 1992

Date

Jack A. Kavanagh

Director's Signature