

# COMPARISON STANDARDS: EXPLORING THEIR VARIETY AND THE CIRCUMSTANCES SURROUNDING THEIR USE

Sarah Fisher Gardial, University of Tennessee  
Robert B. Woodruff, University of Tennessee  
Mary Jane Burns, Florida International University  
David W. Schumann, University of Tennessee  
Scott Clemons, University of Indiana-Southeast

## ABSTRACT

Satisfaction research has made a significant contribution to the field of consumer behavior by exploring the role that comparison standards play when consumers make product-related judgments. However, little empirical attention has been given to documenting 1) the different types of standards which consumers might use or 2) how the evaluation context might influence the consumer's selection of different comparison standards. In this study, in-depth interviews were conducted with members of a health club during which they were asked to discuss their use experiences (comparison standards were neither cued nor prompted). From these interviews, a protocol analysis was conducted which verified a variety of unprompted comparison standards which the consumers used when discussing the service faculty. In addition, the comparison standards which consumers used varied with the means-end hierarchy level at which the judgment was being made, and to a lesser extent with judgments which resulted in emotion versus evaluation responses. Implications for theory development, measurement, and future research are noted.

## INTRODUCTION

Satisfaction research has made a significant contribution to the field of consumer behavior by exploring the role that comparison standards play when consumers make product-related judgments. Both theory and empirical evidence leave little doubt that consumers use standards in these evaluations. For example, disconfirmation theory and supporting research have shown that expectations-based disconfirmation helps explain variation in satisfaction responses (e.g., Oliver 1980). Even in those cases where perceived performance alone seems to be best at explaining satisfaction (e.g., Churchill and Surprenant 1982), it may be that consumers are thinking of some

standard when rating a product's performance. That is, performance measures themselves may be implicitly comparative in nature.

Because consumers use comparison standards in evaluation judgments, these standards say a lot about the nature of competition that a company faces in the market place. Typically, we think of competition as rivalry between companies for purchase choice. A standard may reflect this kind of competition when consumers make product judgments by comparing the focal brand to the performance of another brand. This could happen both before and after product consumption. Both choice (e.g., Biehal and Chakravarti 1993) and satisfaction research (e.g., Woodruff et al 1991) support this notion.

However, use of standards raises the possibility that competition may have a meaning outside of rivalry. The perceived performance of one product may affect evaluation judgments of another even if a consumer did not choose between them. For example, the authors are aware of a company's senior marketing manager who believes that consumers of its sport cruiser boats expect performance, particularly with regard to reliability, to live up to that of cars. Feedback from the company's customers support this view. Though these customers are not choosing between cars and boats, they are using the performance of cars to judge satisfaction with their boats. In this sense, sport cruiser boats are "competing" with cars in the minds of these consumers.

In spite of the important role of comparison standards, and their frequent mention in choice and satisfaction literatures, there remain critical issues to address. For one thing, there have been many standards conceptualized, but only a few have received much empirical attention. For instance, predictive expectations as a standard has been studied extensively, and equity has a growing body of support (Oliver and Swan 1989). Other standards have not been operationalized or examined nearly as well. This raises an important

issue: (1) what are the different kinds of comparison standards that consumers actually use when making product judgments?

If consumers use more than one kind of standard, and there are indications that this is the case (e.g., Cadotte, Woodruff and Jenkins 1987), then a second issue should be examined: (2) are there differences in the circumstances when some standards might be used more than others? For example, time might be an influential circumstance. That is, consumers might be more likely to use one kind of standard soon after purchase, while other standards become more likely at times further removed from the purchase act (Gardial et al 1993). Addressing this issue requires conceptualizing the kinds of circumstances of interest and then looking for the standards being used.

The purpose of this study is to address these two general issues. As part of a larger study, data were gathered to explore the use of comparison standards for evaluation judgments made during the use of a service. Before discussing this research, we briefly review satisfaction literature's contribution to comparison standard thought.

### CONCEPTUAL DEVELOPMENT

Conceptually, one can think of a product evaluation as being absolute or comparative. An absolute evaluation would require that consumers be able to rate a product or its performance without reference to something else. For instance, a consumer might evaluate a restaurant waiter's service as fast or slow by assigning these judgments to time in minutes (e.g., slow food service is anything that takes longer than five minutes).

In contrast, a comparative product evaluation would require that the consumer consider something other than the performance itself. This "something else" is a standard against which the performance can be compared. For instance, a the speed of service in one restaurant may be judged relative to another frequently visited restaurant or relative to a perceived "average time" across a set of restaurants. The satisfaction literature has taken the position that product judgments regarding use experiences are comparative in nature. This is reflected in the major satisfaction theories, such as

the disconfirmation paradigm and equity theory.

### The Nature of Comparison Standards

It has been shown that satisfaction responses are sensitive to the specification of a comparison standard to frame satisfaction questions (e.g., Swan and Trawick 1979; Cadotte, Woodruff and Jenkins 1987; Oliver and Swan 1989). At least one author has suggested that this characteristic makes satisfaction data of questionable usefulness for policy makers (Olander 1977). However, this conclusion is premature. Suppose researchers could determine which standard was being used by consumers under specified circumstances. Then, satisfaction data might become even more useful, particularly for diagnosis of product strengths and weaknesses. Unfortunately, we do not have sufficient research evidence available to say whether or how precisely researchers can pinpoint comparison standards being used by consumers in a target market.

To more fully assess the effect of standards on satisfaction, we need to learn about the different types of standards that consumers might use. The satisfaction literature offers some help on this matter. Conceptual categories have been discussed by several authors (Woodruff et al 1991; Tse and Wilton 1988), and so it is not necessary to review different types of standards here. In fact, this study was based on the categories in Woodruff et al (1991). However, satisfaction literature offers very little beyond these classifications that can help to address the research issues raised above.

### Frequency of Standards Use

Only one study could be found that examines the frequency of use of different types of standards. Gardial et al (1993) considered five types of standards -- product category, product type, other brand, same brand, internal and marketer. They found that all were used by consumers in evaluation judgments, but that some were much more frequently used than others. "Other brand" was by far the most frequently mentioned comparison standard, followed by "internal standards" and "same brand" in decreasing order. Surprisingly, the other standards evidenced relatively low frequency of use. This

---

study suggests an important research question on which further evidence is needed:

Q1: Do consumers use multiple types of comparison standards during postpurchase product (or service) evaluations and, if so, with what frequency?

### Circumstances for Use of Standards

A growing number of studies have examined the circumstances or contexts under which different standards are used. Churchill and Surprenant (1982) found that use of a standard varied by type of product. Expectations disconfirmation influenced satisfaction with a nondurable product (flowers), but not for a durable product (video disk player). Similarly, Bolting and Woodruff (1988) showed that situational involvement may affect the type of standard applied to product performance evaluations. Specifically, in low involvement situations consumers were more likely to use a favorite brand norm or focal brand expectations than a product norm in disconfirmation judgments.

In a study that compared pre- and postpurchase product evaluations, Gardial et al (1993) discovered that frequency of standards used differed across the purchase/consumption process. For example, "other brands" as a standard was used more than twice as much in postpurchase evaluations than in prepurchase evaluations. Finally, the work that applies equity theory to satisfaction suggests that for interactions with sellers, consumers may use equity as a kind of standard for evaluating performance experiences (Oliver and Swan 1989).

All of these studies support the general notion that the context in which product evaluation judgments are made will influence the type of standard applied. However, research has only scratched the surface of possible contexts. One particular context of interest to the authors is the level in a means-end hierarchy at which consumers are thinking. There is growing evidence to support the notion that consumers hierarchically link product attributes to consequences and consequences to purposes or values (Gutman 1982; Reynolds and Jamieson 1985). If so, consumers may use different types of standards to evaluate

product performance at different levels, such as on attributes versus consequences. This suggests the following question:

Q2: Do consumers use different types of standards when they are thinking at different levels within a means-end hierarchy?

There has been considerable interest in the satisfaction literature over the cognitive versus emotional content of satisfaction responses (Westbrook and Oliver 1991, Gardial et al 1993). Some evaluations may be devoid of much feeling, while others are more emotional in nature. One factor that may be associated with the amount of emotional content is the type of comparison standard used to evaluate product performance. Thus, another research question is suggested:

Q3: Do consumers use different types of standards in product evaluations that differ in emotional content?

Finally, there has been some speculation about whether satisfaction and dissatisfaction are opposite ends of a single continuum or two different phenomenon. To date, the evidence is inconclusive (Woodruff et al 1992). However, one way to explore this issue is to determine if different antecedents are associated with satisfaction versus dissatisfaction. For instance, if different comparison standards are applied to product performance evaluations leading to satisfaction than to dissatisfaction, then the two might be better conceptualized as different phenomenon. This suggests the following research question:

Q4: Do consumers use different comparison standards when making product evaluations that lead to satisfaction than they do when dissatisfaction is the outcome?

The purpose of this paper is to present and discuss the results of research that examined these four research questions.

## METHODS

### The Study Design

Most postpurchase and satisfaction research is based upon one-shot measurements, despite the fact that many consumer products are used and consumed repeatedly after choice, providing multiple opportunities for evaluation. For this reason, one of the primary objectives of this research was to engage in a longitudinal study which would capture consumers' postpurchase evaluation processes at various points after choice. Specifically, the respondents were first interviewed within their initial month of product use, and subsequent follow-up interviews were conducted between seven and ten months later. For the purposes of this paper, only the data from the initial interview will be analyzed. Later papers will address the more complex issue of the extent to which evaluation processes might differ at later points in the consumption cycle.

The "product" which was chosen, a health and fitness center, was selected because of its unique combination of physical offerings (equipment, facilities, etc.) as well as a significant service component (staff, instructors, training and fitness evaluators, etc.). In generalizing from this type of product, it should also be kept in mind that this is a fairly complex product (a multitude of features and benefits which can potentially be evaluated) as well as highly involving for many consumers relative to many other consumer product purchases. This involvement is most likely a combination of monetary and psychological considerations in the choice and use of a fitness center.

### Sample

Respondents were identified by the health and fitness center management as individuals in their first month of membership. While twenty-two individuals were identified, one respondent moved away prior to the second interview and a second did not meet our criteria for inclusion in the study (e.g., this person had joined but not used the club), leaving a sample size of 20 for data analysis. These individuals represented a fairly broad demographic profile, although it was skewed

towards a higher socioeconomic profile: average age was 44 (range=25-70); 53% were female; the average income was \$50,000 (range=less than \$15,000 to greater than \$100,000); all but one respondent had at least some college education, with ten having post-graduate education; the average household size was 2.9 individuals (range=1-5); and eleven respondents' occupations were categorized as professional (including health care, sales, engineering, teaching), while other occupations included homemakers, skilled workers, and one retiree.

### The Interview Process

Respondents for the study were first contacted by letter to solicit their participation in the study. Using letterhead from a local university and introducing a team of faculty researchers, respondents were told that the objective of the study was to find out about consumers' health club experiences, that they would be interviewed on 2-3 occasions, and that they would be paid for their participation. A follow up contact was then made by telephone at which time the interviewer confirmed the respondent's agreement to participate and set up the initial interview. At this time, the interviewer also obtained the respondent's permission to video tape the interview and assured him/her of confidentiality.

All interviews were held at the health and fitness center and were conducted by a graduate student who had been trained in depth interviewing techniques. An interview room was set up to provide privacy as well as accommodate the videotaping equipment. Subjects were, again, asked for their permission to tape and reassured of confidentiality. The interviewer then conducted a depth interview to probe respondents' descriptions, reactions, evaluations, and feelings with respect to their use of the health and fitness center.

The format of the interview was semi-structured. A series of common questions was asked across subjects, but the interviewer was free to probe responses and follow up on respondent's unique concerns and observations. The questions were open-ended and were constructed to provide ample opportunity for respondents to describe their experiences at the center. (See appendix A for a list of the interview questions.) At no time in the

interview were the respondents explicitly asked about comparison standards, or whether or how they used them.

### Data Coding

The respondents' interviews were transcribed for use in a protocol analysis. These transcripts were broken into "thoughts," each thought representing a distinct idea which was conveyed by the respondent. Because of differences in verbalization, the actual length of thoughts varied across and within respondents. Two researchers independently circumscribed the thoughts within each transcript and then discussed, reconciled, and agreed upon what constituted the thoughts in each transcript. A total of 7463 thoughts were identified across the twenty transcripts, resulting in an average of 373 thoughts per subject (range=275-476).

A coding scheme was designed to identify and define processing characteristics of interest, both for this study and others which will not be considered here. This coding scheme was used to classify the content of the transcript thoughts. With respect to this study, categories and definitions were created to help identify 1) different types of comparison standards which might appear in the thoughts, 2) references to various levels within the means-end hierarchy, 3) the presence of either evaluation outcomes or emotion responses, and 4) thoughts which included specific mention of either satisfaction or dissatisfaction. (For more detail, see Appendix B.)

It was possible for each thought to receive four codes, one for each of the categories above, although a "not mentioned" response was available in each category. For the purposes of this paper, the data set includes the 1378 thoughts which contained a comparison standard (a positive response to category 1 above). Beyond this, each "standard" thought varied in whether it could also be classified into one or more of the remaining categories (categories 2-4 above). Therefore, the number of codes per thought varied.

Two graduate students, who were independent of the research project, coded the data. They were trained in coding techniques, familiarized with the coding categories, and periodically evaluated

throughout the coding process to monitor consistency and thoroughness. Each coded the thoughts independently, and then the two coders reconciled any disagreements. Disagreements which could not be reconciled were resolved by one of the researchers. These coders had a 92 percent agreement rate.

**Table 1**  
Respondents' Mean Number of Thoughts  
Per Coding Category

<u>Standards</u>							
<u>Total</u>	<u>Attri- butes</u>	<u>Conse- quences</u>	<u>Values</u>	<u>Evalu- ations</u>	<u>Emo- tions</u>	<u>Sat.</u>	<u>Diss.</u>
<u>Product</u>							
22.65	12.85	5.25	2.40	9.10	0.60	0.30	0.05
<u>Other People</u>							
8.85	2.40	1.95	1.45	1.20	0.35	0.45	0
<u>Other Sitn.</u>							
17.90	4.10	11.95	5.35	4.00	1.30	0.60	0.05
<u>Other Time</u>							
7.75	0.80	2.05	0.85	1.00	0.25	0	0
<u>Internal</u>							
2.20	0.95	0.50	1.05	1.00	0.15	0.15	0
<u>Marketer Supp.</u>							
0.40	0.30	0.05	0.10	0.15	0	0.05	0.10
<u>Other</u>							
9.15	3.85	3.10	2.20	2.10	0.80	0.40	0.15
<u>Total</u>							
68.90	25.25	24.85	13.40	18.55	3.45	1.95	0.35

### Data Analysis

The relevant data for this study are included in Table 1. These are the average number of responses per respondent which fell into the respective coding categories. A paired comparison t test was used to test differences within the data set, and a more conservative p value ( $p < 0.025$ ) was used to compensate for the use of multiple comparisons. This was calculated by dividing the normal cut off ( $p < 0.05$ ) by two, so that the "best" response in each category was compared against the next highest response (see Stephansson, Kim and Hsu 1988).

## RESULTS

### Research Question One

The first research question was whether respondents would use a variety of standards of comparison in their responses to product use. This was found to be the case. Respondents in this study used all six of the specific types of standards defined in the coding scheme, as well as miscellaneous other standards. Products were most often used as the comparison (mean=22.65), followed by other use experiences (mean=17.9), other people (mean=8.85), and other points in time (mean=7.75). Less frequently used were the respondent's internal standards (mean=2.20) and marketer supplied expectations (mean=0.40). "Other" standards represented approximately 13% of all standards used. It is interesting to note that no one type of standard clearly dominated, with the largest single response category (product standards) representing slightly less than 1/3 of all standards used.

Also important to note is the extent to which an individual's consideration of other product use experiences were used to shape standards; other situations and other points in time combined to represent some 37% of the standards used. Finally, although marketer supplied expectations were not found as frequently, this may be due to the fact that these respondent's had had relatively little time to use the product and either verify or falsify sales claims. It is possible that we will see an increase the incidence of these standards in the subsequent interviews.

In sum, respondents were able to use a variety of comparison standards in responding to their product use experiences. This is consistent with previous research results, although the health and fitness center represents a unique product category.

### Research Question Two

Research question two asks whether the types of comparison standards used might vary according to the level in the means-end hierarchy. We found evidence that standard do differ by level. In particular, as the respondent moves up the hierarchy it is important to note the shift from product-based standards toward standards which

relate to respondent's prior product use experiences. At the attribute level, not surprisingly, other products (mean=12.85) are the dominant comparison standard, being used significantly more often than the next most frequent category, other situations (mean=4.10,  $p < 0.0002$ ). Additional standards used to evaluate attributes included other people (mean=2.40), internal standards (mean=0.95), other points in time (mean=0.80), marketer supplied expectations (0.30), and other (mean=3.85).

However, at the consequence level, "other situations" were used more often than products (11.95 versus 5.25, respectively), and the same is true at the value level of the hierarchy (5.35 versus 2.40, respectively). Combining the data for the higher means-end levels, we find that "other situations" were used significantly more than "product" standards (17.30 versus 7.65,  $p < 0.013$ ).

Less frequently used standards when considering product consequences were other points in time (mean=2.05), other people (mean=1.95), internal standards (mean=0.50), "other" standards (mean=3.10), and marketer supplied standards (mean=0.05). Less frequently mentioned standards for values included other people (mean=1.45), internal standards (mean=1.05), other points in time (mean=0.85), marketer supplied standards (mean=0.10), and "other" standards (mean=2.20).

In sum, it appears that subjects use different comparison standards at different levels of the means-end hierarchy, with product standards prevalent at lower means-end levels and an increasing emphasis on standard related to product use experiences at the higher levels.

### Research Question Three

The third research question is concerned with whether different comparison standards might be associated with evaluation versus emotion outcomes. Again, we find evidence to suggest that this might be true. "Product" standards clearly dominate the comparison standards associated with evaluation outcomes (mean=9.10 or approximately half of all standards), while "Other Situations" are used significantly less often (mean=4.00,  $p < 0.0051$ ). In contrast, emotion outcomes tend to be

more strongly associated with product use situations; "other situation" standards are mentioned twice as often (mean=1.30) as product standards (mean=0.60,  $p < 0.1625$ ). Although this result is approaching significance, the test is constrained by the small sample size due to the infrequent mention of emotion outcomes.

Less frequently mentioned standards associated with evaluations were other people (mean=1.20), other time (mean=1.00), internal standards (mean=1.00), marketer supplied standards (mean=0.15), and "other" standards (mean=2.10). Less frequently mentioned standards associated with emotion responses were other people (mean=0.35), other time (mean=0.25), internal (mean=0.15), and "other" (mean=0.80).

#### Research Question Four

Question four suggested that there might be differences in the types of comparison standards which were associated with satisfaction versus dissatisfaction. However, respondents rarely mentioned the terms "satisfaction" or "dissatisfaction," and so there were insufficient data available to fully address this question. One reason for this lack of data is probably the relative inexperience that the subjects had with the product prior to this initial interview. They may not have had time to form clear judgments of (dis)satisfaction. Again, we might expect to see an increase in these types of evaluations over time. In addition, these data are consistent with other research which indicate that consumers very rarely use the words "satisfaction" and "dissatisfaction" in describing their product experiences (Woodruff et al. 1992).

This qualification being noted, the preliminary results are intriguing and are consistent with previous research. The use of product standards appears to be more frequently associated with satisfaction than with dissatisfaction evaluations (0.30 versus 0.05, respectively), while marketer supplied expectations tend to be more frequently associated with dissatisfaction than with satisfaction (0.10 versus .05, respectively).

## DISCUSSION

Several limitations of this study should be noted prior to discussing its implications. Specifically, these results may be unique to the specific product category (a health and fitness center), the sample of subjects (fairly high socio-economic status), and their particular stage in the consumption process (in the early stages of consumption). However, the authors' previous research, which has been carried out to date across multiple product categories, has accumulated increasing evidence which supports the generalization of these findings.

The results of this research support previous work suggesting that consumers use multiple comparison standards in product evaluation judgments. Standards used are based on consumers' prior experience, such as performance of products and brands with which they are familiar, as well as comparing product performance across various use situations, times and people. Further, the context in which judgments are made has long been known to be important (Chakravarti and Lynch 1983), and comparison standards do appear to differ by context. These findings have implications for both satisfaction theory and research and for the application of satisfaction data in marketing practice.

#### Satisfaction Theory and Research

**Theory.** The results of this study provide support for satisfaction theory's emphasis on comparative evaluations of products. Importantly, respondents were not specifically asked to list or discuss the comparison standards they might use. Instead, respondents discussed their use experiences with the fitness center service, and in this broader context they chose to mention standards. This suggests that comparison standards are a part of the way in which consumers think about performance evaluations, which is quite consistent with the disconfirmation notion so prevalent in satisfaction theory.

As Woodruff et al (1991) argued, however, disconfirmation theory may be too limited by focusing only on expectations as the standard that consumers use in satisfaction judgments. The

results of this study support their position. Without solicitation, respondents mentioned several different types of comparison standards when discussing their use experiences. Thus, satisfaction theory should be broadened to account for standards that are important to consumers beyond just predictive expectations.

The study suggests that use of standards is context driven. Where the consumer is in the means-end hierarchy and, to a lesser extent, the degree of evaluation versus emotion seem to be associated with type of standard mentioned. This means that satisfaction outcomes are not likely to be fully explained by a simple expectancy-disconfirmation process. Satisfaction theory will have to account for contingencies. In a broad sense, work is already preceding in this direction by introducing what can be viewed as contingency constructs into satisfaction theory, such as attribution (Oliver 1989), equity (Oliver and Swan 1989), emotion (Westbrook and Oliver 1991), and experienced-based norms (Woodruff, Cadotte and Jenkins 1983). This research suggests another direction for building a contingency satisfaction theory. Such theory should incorporate constructs and processes that explain how different types of comparison standards might affect satisfaction outcomes.

**Future Research.** This study was exploratory and it provided no indication of the causal role of standards in satisfaction processes. Before a contingency theory can be fully specified, additional research is needed that relates to the role of comparison standards. Certainly, a starting point would be to develop a classification scheme that would organize standards into categories based on dimensions. For example, dimensions might be related to product, situation, time, and the like. Then, research could be devoted to understanding how various comparisons standards work within satisfaction processes. These general directions suggest several research questions, that if answered, would provide important insights into the directions that theory building should take. These include:

1. Is it possible to develop a classification scheme for comparison standards used on postpurchase evaluations, based on selected

dimensions? For example:

- a. the dimension of time may be one possible underlying dimension; are seller's promises more likely to be applied to evaluations of product performance close to the time of purchase or further removed from purchase?

- b. are certain kinds of standards more likely to be applied by consumers in some product use situations more than others? Could the construct of consumer problem space be another possible dimension?

2. Do satisfaction outcome measures significantly differ when respondents are thinking about product performance in different use situations? Similarly, if satisfaction outcome measures do not specify use situation, do respondents answer in the context of a use situation of their choice or more generally across use situations?

3. Are certain types of standards more likely to lead to satisfaction outcomes for product performance evaluations than to dissatisfaction? For example, is comparison to seller's promises more likely to lead to satisfaction than comparison to an internal standard based on need?

4. Are certain types of standards more likely to lead to satisfaction outcomes that have high emotional content than do other standards?

5. Do standards play a causal role in product performance evaluations, as would be described by a disconfirmation-type process, or might standards be selected to be consistent with satisfaction outcomes?

### **Applications of Satisfaction in Practice**

While it may be premature to speculate about implications of an exploratory study's findings for marketing practice, several thoughts did come to mind. For one thing, if consumers tend to respond differently to disconfirmation questions based on different standards, companies may have to



become more proactive in selecting standards for satisfaction measures. In particular, it may be that by selecting a specific comparison standard in measures, more actionable information would be obtained for decisions. For instance, suppose satisfaction data are to be used to stimulate ideas for advertising campaigns. Satisfaction data where respondents indicate disconfirmation relative to a major competitive brand may yield best insights on the company's brand strengths to emphasize. But for product design, it may also be important to understand how consumers evaluate a product relative to internal desires. In general, different standards may yield disconfirmation or satisfaction data yielding complementary insights into market opportunities.

Several authors have argued that satisfaction judgments may be made at higher levels in a means-end hierarchy than at the attribute level (Olshavsky and Spreng 1989, Clemons and Woodruff 1991; Woodruff et al 1991). To learn about these judgments, say at the consequence level, might require designing satisfaction measures that incorporate standards most likely to be used at that level. For example, the study indicated that situational standards are more likely to be used than product standards when respondents are thinking at a consequence level. Again, a company may have to be proactive in selecting which standards to employ in the data collection effort.

Finally, although the occurrence of emotion in this study was insufficient to make firm recommendations, the study suggests that certain standards may be more likely to be associated with emotional content in postpurchase evaluation judgments. For example, situational standards were more associated with postpurchase evaluations expressed with emotion words than were product standards. To the extent that positive emotion experienced by customers is desirable for a company, some attention must be paid to the comparison standards that are associated with emotion. A company may want to use promotion to communicate comparison standards to customers, making certain ones salient during actual satisfaction judgments.

In sum, the results from this study offer new insights in the types of comparison standards used by consumers in post-purchase evaluation

processes. Some new types were found, such as other people, situations, and time, that have not been previously mentioned in satisfaction literature. Further, the study suggests that comparison standards are related to important postpurchase evaluation phenomena such as the means-end hierarchy. These results should encourage continuing research on the role that comparison standards play in evaluations, including exploring ways to improve the actionability of consumer satisfaction data.

## REFERENCES

- Biehal, Gabriel and Dipankar Chakravarti (1983), "Information Accessibility as a Moderator of Consumer Choice," *Journal of Consumer Research*, 10 (June), 1-14.
- Bolting, Claire P. and Robert B. Woodruff (1988), "Effects of Situational Involvement on Consumers' Use of Standards in Satisfaction/Dissatisfaction Processes," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 1, 16-24.
- Cadotte, Ernest R., Robert B. Woodruff and Roger L. Jenkins (1987), "Expectations Versus Norms in Models of Consumer Satisfaction," *Journal of Marketing Research*, 24 (August), 305-314.
- Chakravarti, Dipankar and John G. Lynch (1983), "A Framework for Exploring Context Effects on Consumer Judgments and Choice," in Richard P. Bagozzi and Alice M. Tybout, eds., *Advances in Consumer Research X*. Ann Arbor: Association for Consumer Research, 289-297.
- Churchill, Gilbert A. and Carol Surprenant (1982), "An Investigation into the Determinants of Customer Satisfaction," *Journal of Marketing Research*, 19 (November), 491-504.
- Clemons, D. Scott and Robert B. Woodruff (1991), "Broadening the View of Consumer (Dis)satisfaction: A Proposed Means-End Disconfirmation Model of CS/D," in Chris T. Allen et al, eds., *Marketing Theory and Applications*. Chicago: American Marketing Association, 1992, 413-421.
- Gardial, Sarah Fisher, D. Scott Clemons, Robert B. Woodruff, David W. Schumann, and Mary Jane Burns (1993), "Comparing Consumers' Recall of Prepurchase and Postpurchase Product Evaluations," *Journal of Consumer Research*, forthcoming.
- Gutman, Jonathon (1982), "A Means-End Chain Model Based on Consumer Categorization Processes," *Journal of Marketing*, 46 (Spring), 66-72.
- Olander, Folke (1977), "Consumer Satisfaction - A Skeptic's View," in H. Keith Hunt, ed., *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*. Cambridge, MA:

- Marketing Science Institute, 72-91.
- Oliver, Richard L. (1980), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, 17 (November), 460-469.
- Oliver, Richard L. (1989), "Processing of the Satisfaction Response in Consumption: A Suggested Framework and Research Propositions," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 2, 1-16.
- Oliver, Richard L. and John E. Swan (1989), "Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions: A Field Survey Approach," *Journal of Marketing*, 53 (April), 21-35.
- Olshavsky, Richard W. and Richard A. Spreng (1989), "A 'Desires As Standard' Model of Consumer Satisfaction," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 2, 49-54.
- Reynolds, Thomas J. and Linda F. Jamieson, "Image Representations: An Analytic Framework," in Jacob Jacoby and Jerry C. Olson, eds., *Perceived Quality: How Consumers View Stores and Merchandise*. Lexington, MA: Lexington Books, 115-138.
- Stephansson, Gunnar, Woo-Chul Kim and Jason C. Hsu (1988), "On Confidence Sets in Multiple Comparisons," *Statistical Decision Theory and Related topics IV*, Vol. 2, Springer-Verlag, New York, Inc., 89-104.
- Swan, John E. and I. Fredrick Trawick (1979), "Satisfaction Related to Predictive vs. Desired Expectations," in H. Keith Hunt and Ralph L. Day, eds., *Refining concepts and Measurements of Consumer Satisfaction and Complaining Behavior*. Bloomington, IN: School of Business, Indiana University, 7-12.
- Tse, David K. and Peter C. Wilton (1988), "Models of Consumer Satisfaction Formation: An Extension," *Journal of Marketing Research*, 25 (May), 204-212.
- Westbrook, Robert A. and Richard L. Oliver (1991), "The Dimensionality of Consumption Emotion Patterns and Consumer Satisfaction," *Journal of Consumer Research*, 18 (June), 84-91.
- Woodruff, Robert B., Ernest R. Cadotte and Roger L. Jenkins (1983), "Modeling Consumer Satisfaction Processes Using Experience-Based Norms," *Journal of Marketing Research*, 20 (August), 296-304.
- Woodruff, Robert B., D. Scott Clemons, David W. Schumann, Sarah F. Gardial, and Mary Jane Burns (1991), "The Standards Issue in CS/D Research: A Historical Perspective," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 4, 103-109.
- Woodruff, Robert B., David W. Schumann, D. Scott Clemons, Mary Jane Burns, and Sarah F. Gardial (1992), "The Meaning of Satisfaction and Dissatisfaction: A Themes Analysis from the Consumer's Perspective," working paper, The

University of Tennessee.

### Appendix A Interview Questions

1. What do you think about the health and fitness boom that is currently going on?
2. What experiences have you had with other clubs?
3. How did you come to join (this particular) club?
4. Can you start at the beginning and take me through a typical visit to the club - step by step from the time you leave your car in the parking lot.
5. Do any of your visits here stand out?
6. If you were going to tell me what this experience meant to you, what words would you use? Suppose someone asked you what you thought of the club, what would you say?
7. Are you satisfied with this membership?
8. Are there any aspects of this club that you are dissatisfied with?
9. Do you value your membership? Does your membership have value?
10. What does your membership do for you?

### Appendix B Coding Categories and Examples

**Type of Standard Used.** This category was used if the respondent compared the product or his/her use of the product with a comparison standard. Seven possible types of standards were identified.

1. **Product.** Another product was used as a point of comparison. This could either be in the same product category (e.g., another health and fitness center), from a subgroup within the product category (e.g., a swimming club), or from a completely different product category (e.g., participation in a sports league). For example, "I find this club to be much cleaner than (my previous health club)."
2. **Other People.** The respondent's comparison standard is with another person, i.e., how their experience with the product might differ from another person's experience. For example, "I just don't get as much out of this club as people who work out here on a daily basis."
3. **Other Situations.** The comparison standard is other situations which the respondent has experienced, i.e., s/he compares product performance across use situations. For example, "I enjoy my workouts more when there are others around than when I am by myself."
4. **Other Time.** The respondent's comparison standard is to another point in time. For example, "I didn't work out nearly as often when I was working the night shift."

5. **Internal.** The respondent's comparison standard is to some internalized ideal or "want" that the consumer has in mind. For example, "It's very important to me that the locker rooms be immaculate."

6. **Marketer Supplied.** The respondent's comparison standard is a promise or assurance about the product that was provided by the marketer or salesperson. For example, "They told me that there would be free baby sitting."

7. **Other.** A miscellaneous category for responses that did not fit the above categories.

**Send correspondence regarding this article to:**

Sarah Fisher Gardial

Department of Marketing and Transportation

University of Tennessee

Knoxville, TN 37996-0530 USA

**Means-End Hierarchy Levels.** This category was used when the respondent's thought also included reference to a level within the means-end hierarchy.

1. **Attribute.** The product characteristic mentioned is at the attribute or feature level. For example, "I like the equipment."

2. **Consequence.** The product characteristic mentioned is at the consequence level, i.e., it is either a benefit or a sacrifice that the respondent experiences during product use. For example, "I got a really good workout without getting sore."

3. **Values.** The subject mentions either a personal, organizational or role value which is attached to product use. For example, "The family plan allows me to spend more time with my children."

**Outcomes.** These categories are used to classify the outcomes of evaluation processes.

1. **Evaluation Outcomes.** The respondent's processing results in a good-bad response toward the product. For example, "I loved the aerobics instructor."

2. **Emotion Response.** The respondent's evaluation results in a specific emotion. These emotions had to be explicitly stated in the transcript (could not be inferred by the coders) and a list of emotion words was used to aid in their identification. For example, "I was really angry when they changed the fee on me."

**Satisfaction.** These categories were used when the respondent explicitly referred to (dis)satisfaction in responding to his/her use of the product, either unprompted or when prompted by the interviewer.

1. **Satisfaction.** The respondent explicitly uses the word "satisfaction" in response to product use.

2. **Dissatisfaction.** The respondent explicitly uses the word "dissatisfaction" in response to product use.