EMOTIONAL MEASURES OF SATISFACTION/DISSATISFACTION

Douglas Hausknecht, University of Akron

ABSTRACT

This paper utilizes the critical incident technique and the Differential Emotions Scale to derive emotion-based definitions and measures of satisfaction and dissatisfaction. A preliminary investigation using student subjects concluded that satisfaction is best characterized by the emotions interest, joy and surprise; dissatisfaction by anger, disgust and surprise.

INTRODUCTION

The marketing literature in general and the consumer satisfaction/dissatisfaction (CS/D) literature in particular have been hindered by conflicts in measurement development and use. Part of this has resulted from incomplete or inconsistent construct definitions.

Interest in satisfaction was rooted in the desire to understand the motivation underlying post-purchase behaviors,

> Customer satisfaction with a product presumably leads to repeat purchases, acceptance of other products in the same product line, and favorable word-of-mouth publicity. If this assumption is correct, then knowledge about factors affecting customer satisfaction is essential to marketers.
>
> Cardozo 1965, p. 244

More recently, however, the CS/D literature appears to have bogged down in cognitive measures of attitude-like phenomena.

A notable exception is Westbrook's (1983) investigation of post-purchase emotions. Although not proposed as a new satisfaction measurement technology per se, the emotion scale used may be adaptable for this purpose. If this can be accomplished, it would be the first satisfaction scale which encompasses the critical motivational components in a format that can be applied to a wide range of products and situations. (See Singh and Howell 1985 for a related discussion focussing on the resultant complaint behaviors.) This discussion, then, is consistent with the definition of satisfaction proposed by Day, "...an emotional response manifested in feelings and is conceptually distinct from cognitive responses, brand affect and behavioral responses" (1983, p. 113).

Present Satisfaction Measures

A wide variety of measurement instruments have been proposed and utilized in consumer satisfaction studies. Westbrook and Oliver (1981; Oliver and Westbrook 1982) have compared a number of the measures which have been developed. From two analyses of pilot data they concluded: 1) Likert, semantic differential and a composite verbal scale (see Table 1) performed best on convergence versus divergence criteria; 2) discriminability of various scales seemed to be product class dependent (e.g., automobiles versus calculators); and 3) as a whole, the measures did not succeed very well in discriminating satisfaction from attitude. The first finding is encouraging for those who wish to identify measures which will be useful across situations. The latter two findings threaten the universality of both the measures and theories of satisfaction.

The conclusions by Westbrook and Oliver are tempered greatly by their methodology and the ensuing discussion. Even given that the student sample (n = 160) is sufficiently representative for this type of research (theory testing), the manner of questioning still leaves room for criticism. Essentially, the students were asked to give their present reaction to durables which they may have possessed for some time and with which they are likely to have varying histories of ownership and usage experience. It is not surprising, then, that the measures reflect some difference in response to product class and some confounding with attitude. The authors allude to an explanation "...satisfaction...is, in effect, a response to disconfirmation... and is expressed in affective terms... In a sense, satisfaction may be seen as a disturbance acting on an attitude system." (1982, p. 13). They go on to suggest that temporally distinct measures may help to separate satisfaction and attitude. Given Oliver's (1981) argument that satisfaction is likely to rather rapidly "decay" into attitude, it would seem that temporally distinct methodology would be necessary to uniquely identify the constructs.

In the absence of such temporal distinction, the cautionary conclusions by Westbrook and Oliver are suspect. This treatment of the theory relating satisfaction and attitude is continued in research which attempts to calibrate emotional measures derived from the Differential Emotions Scale. In this research (Westbrook 1983) the methodology comprised of collecting contemporaneous reactions to products (i.e., automobiles) which have been owned and used over a range of time periods and usage patterns is repeated. The resulting emotional dimensions are correlated with more typical satisfaction measures with mixed results. Again, the theoretical distinction between satisfaction and attitude is not maintained and the influence of time is disregarded. The measurement technique proposed here is used in conjunction with the critical incident methodology in the present paper to alleviate this time problem in the derivation of measures.

Similar concerns about timing versus attitude formation have also been expressed outside the "mainstream" CS/D literature (Locker and Dunt 1978). In this review of British studies concerning the measurement of patient satisfaction with medical care, the authors identify many measurement techniques similar to those commonly used in CS/D. They are also concerned about the necessity of developing universal measures of satisfaction to allow for inter-study comparability. The interview technique seems to be much more prevalent in this literature, and the findings suggest that global evaluations exposed by direct measurement scales seriously mask the separate underlying attribute evaluations.

This argument raises the issue of the specificity of the measures. In order to be useful in theory testing, measures must be applicable at the same level of specificity as the operational theory. Locker and Dunt (1978) argue for examining various aspects of the patient's experience with medical care in order to explain his satisfaction; much the same as Wanous and

Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior, Vol. 1, 1988, pp. 25-33.

TABLE 1
Selected Measures of Consumer Satisfaction/Dissatisfaction

I. Evaluative (Cognitive) Measures

a. Overall, how satisfied have you been with this _____?

100% 90 80 70 60 50 40 30 20 10 0%
Completely Satisfied — Half & Half — Not at all satisfied

b. How satisfied were you with _____?

Very Dissatisfied | Somewhat Dissatisfied | Slightly Dissatisfied | Neither | Slightly Satisfied | Somewhat Satisfied | Very Satisfied

II. Emotional Measures

a. Mark on one of the nine blanks below the position which most closely reflects your satisfaction with _____.

Delighted _____ Pleased _____ Mostly Satisfied _____ Mixed _____

Mostly Dissatisfied _____ Unhappy _____ Terrible _____ Neutral _____

Never Thought About It _____

b. Based on Oliver 1980: (Agree/Disagree)

I am satisfied with _____.
If I had it to do all over again, I would _____.
My choice to _____ was a good one.
I feel bad about my decision concerning _____.
I think that I did the right thing when I decided _____.
I am not happy that I did what I did about _____.

III. Behavioral Measures

a. Complaint/Compliment count

b. How likely are you to use _____ in the future?

Very Unlikely -2 | Unlikely -1 | Likely +1 | Very Likely +2

Lawler (1972) examine the influence of various aspects of a job on job satisfaction. While these are likely to be useful for understanding which factors influence satisfaction (maybe in a specific situation), it is less likely that a respondent will be able to discriminate the separate influences on an inventory of emotions such as that presented by Westbrook (1983) in order to build a "net" satisfaction. The final result of this may be that emotional measures will best reveal the presence and degree of satisfaction whereas evaluative measures (Table 1) will be more useful for explaining the cause of the feeling (i.e., it may be difficult to make emotional measures salient at more than one level in a given measurement situation).

Even so, these measures should not be used without additional cautions. Miller (1977) and others have argued that attempts to measure a process may actually impact upon the process being studied. This problem of reactive measures and concomitant demand artifacts has been discussed elsewhere in consumer behavior (Sawyer 1975) but the arguments do not seem to have had much influence in the design of less obtrusive measures (such as proposed by Webb, Campbell, Schwartz and Sechrest 1966). Sobel and McGuire (1977) have gone so far as to argue that the disconfirmation paradigm (discussed here in a slightly different context) itself may be simply a measurement artifact.

The possibility of this, given current operationalizations of the satisfaction theory should not be overly discounted. Some authors have reported the necessity of forcing a satisfaction response by eliminating the opportunity for neutrality (e.g., Day and Bodur 1979; Jordan and Leigh 1984). While this may be viewed as forcing more careful responses, such a procedure may obscure important theoretical differences between a mild response and no response. For our purposes, it is important to note that the treatment of neutral options such as "Don't Know" or "Never Thought About It" has been shown to bias recall data (Mizerski, Freiden and Greene 1983) and can be presumed to have similar effects on attitudinal or satisfaction responses.

For these reasons, observational measures such as those suggested in Table 1 may actually be better for some purposes than the more direct satisfaction measures identified elsewhere in the Table. Although these behavioral measures may be influenced by other external factors, the causal flow may be sufficiently strong to drive discernible, theoretically important differences. The influence of satisfaction on such behaviors as word-of-mouth, complaining and repurchase is examined in detail in these separate literatures.

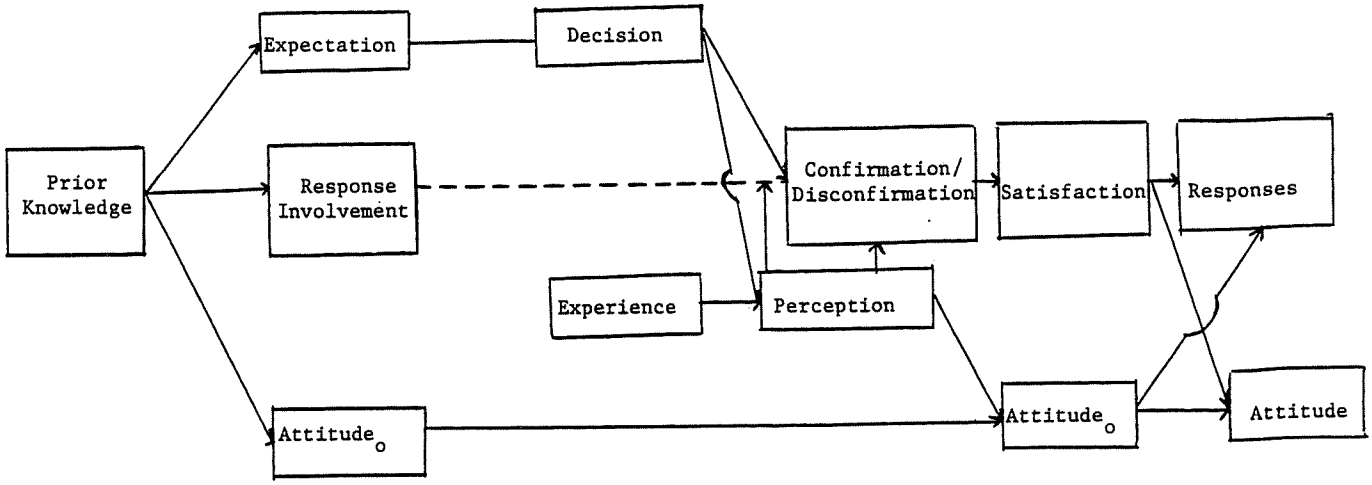
EXPERIMENTAL METHOD

This paper reports a preliminary effort at deriving an instrument to measure satisfaction and dissatisfaction in a manner consistent with the above-described emotional conceptualization. The emotional measures which are derived constitute an operational definition of the construct illustrated in the Figure. The behavioral intention scales, as derived, are also considered to be measures of satisfaction, albeit indirect since they actually assess theoretical outcomes.

Westbrook's studies discussed above provide no criteria by which to identify the emotional profiles which best describe satisfaction. The present study corrects for this in two ways. 1) The critical incident technique which has been introduced to the CS/D literature (Maddox 1981; Swan and Combs 1976) is used to concentrate responses at the ends (negative and positive) of the satisfaction continuum. 2) Measures of intent to engage in specific behaviors thought to result from satisfaction are collected in addition to the more typical satisfaction scales used by Westbrook.

A concise framework for the construction of psychometric scales has been presented in the marketing literature and can be adapted for use in the present study (Lundstrom and Lamont 1976). The principles and philosophy implicit in the scale derivation are consistent with those of the author and those presented in other recent discussions of measure development (Churchill 1979; Nunnally 1978). The Lundstrom and Lamont framework presents five major considerations for the construction of a reliable and valid scale: 1) definition of the construct and selection of an item pool; 2) selection of a scaling procedure; 3) selection of scale items; 4) evaluation of the reliability of the scale; and 5) evaluation of the validity of the scale.

FIGURE
A Model of a General Satisfaction Construct



Based on Day (1983)

The present study examines two methods of measuring satisfaction which has been defined previously as "an emotional response...conceptually distinct from... behavior responses" (Day 1983). One set of measures taps directly the emotional nature of the construct. The second set of measures discussed examines intent to behave in a particular manner. Although this latter set of scales, by definition, can measure the construct only indirectly, the model presented previously indicated the importance of behavioral response to the overall conceptualization (Figure).

Objectives

The principle goals of this study were to identify those emotional measures and those measures of behavioral intent which best discriminate between subjects responding to satisfying and dissatisfying occasions, respectively. Satisfaction scales and criteria for evaluation can then be constructed from these measures.

Method

Ninety-four students in an introductory Marketing class served as volunteer subjects for extra class credit. The study was conducted in various sized groups in a classroom setting and required about 30 minutes per session.

Subjects were given a blank sheet of paper on which to first identify a recent consumer experience which was particularly dissatisfying and then, after all had completed this, identify a recent consumer experience which was particularly satisfying. (This method was chosen because it was felt that a satisfying occasion would be more readily recalled if "primed" by an already salient dissatisfying occasion.) After sufficient time for all subjects to identify an occasion of each type, they were instructed to turn to and complete the first page of the questionnaire. On this page, they were to provide more details concerning either the satisfying (N = 46) or dissatisfying (N = 48) occasion. (In the course of analysis, it was determined that four subjects in the satisfying condition actually described, and completed the remainder of the study responding to, the dissatisfying occasion. This problem and isolated incidences of item non-response created some variations in sample size, as reported below.)

In order to avoid hastily completed and unnecessarily brief descriptions of the specific occasions, subjects were instructed not to continue through the questionnaire until sufficient time had elapsed for a complete narrative (a pre-test indicated that 10 minutes was an adequate period). The remainder of the questionnaire was self-paced and consisted of emotion scales, satisfaction measures and behavioral scales.

Measurement Scales

Westbrook employed an abbreviated version of Izard's Differential Emotions Scale, "...owing to the length of the original scale (69 items)" (1983, p. 5). Yet, he chose not to use the abbreviated version which had already been supplied by Izard (1977). Due to the lack of rational criteria by which to select between these and, in face of Izard's estimate of 5 minutes as the time requirement for a 67-item edition of the scale (1972, p. 86), the full version was administered. In addition, adjectives were extracted from satisfaction scales to round out a 75-item list (added were: Dissatisfied, Satisfied, Pleased, Contented, Thrilled, Depressed, Wise, Foolish). A single randomized order was reproduced on all questionnaires, but individual subjects were assigned separate starting points and instructed to complete the entire list.

Four satisfaction scales were selected to represent common measurement techniques. The scales were chosen on the basis of reliability in prior testing (Oliver and Westbrook 1982; Westbrook and Oliver 1981) and production feasibility (i.e., verbal rather than graphic scales).

The questionnaire concluded with a series of Likert type scales. A form of Oliver's (1980) six-item satisfaction scale began the list and was followed by a series of behavioral intent scales (phrased as behavioral likelihood "I would..." "Strongly Agree - Strongly Disagree"). This latter set was derived from a variety of sources including satisfaction surveys, guides for dissatisfied consumers and personal experience. The behavioral intent measures were intended to be exploratory in nature.

ANALYSIS

The first of Lundstrom and Lamont's (1976) considerations has been satisfied; definition of the construct and identification of the item pool(s). The selection of Izard's DES as the basis for emotional measures identified the scaling procedure as well as the item pool. Thus, a Likert-like scale will be used to measure emotions. For reasons similar to those elaborated by Lundstrom and Lamont (feasibility, simplicity) a Likert (1932) scaling procedure was also selected for the behavioral intent items. Comparisons of Likert and other procedures have generally concluded that the scales which result from the former are at least as reliable (e.g., Thurstone's technique, Edwards and Kenney 1946).

The major task of the Pilot Study involved the selection of specific items as measures of the satisfaction construct. Several analyses of the emotion scales were conducted in order to identify those emotions which best characterize each end of the satisfaction continuum. The 75 items that constituted question 4 were subdivided four ways: items used by Westbrook (1983); items used by Izard (1977) in a shortened scale; those items which exhibited the highest factor scores in Izard's factor analysis; and the full scale.

In general, the analyses conducted were of the following types:

Factor Analysis of Emotions - A varimax principle components factor analysis was conducted for each of the four sets of emotion measures. These analyses were performed to ascertain whether the adjectives said to describe separate emotions would load appropriately. Also, given the restricted emotional patterns expected (i.e., only satisfaction and dissatisfaction), this analysis was expected to suggest which emotions might combine in each pattern.

Discriminant, Cluster and Regression Analysis of Emotions - Scales were created by summing the items loading on each factor resulting from the prior analyses (Table 2). These intermediate scales were then employed in efforts to segment subjects. Discriminant analysis using assigned condition (satisfying versus dissatisfying occasion described), K-means cluster analysis looking for "good" two-cluster solutions and regression of more standard satisfaction measures on the emotional scales were expected to result in identification of the emotions most representative of satisfaction and dissatisfaction.

Factor Analysis and Other Analyses of Behavioral Intent Measures - Rather than attempt to measure and scale actual behaviors in the Pilot Study, Likert scales of potential satisfaction related behaviors were used as a surrogate. As with the emotional scales a varimax principle component factor analysis was used to derive intermediate scales which were used in further analyses (Table 4). Discriminant analysis using assigned condition and mean comparisons across groups defined by emotional profiles were conducted to identify measures of behaviors most typical of satisfied and dissatisfied respondents.

Results

Given the exploratory nature of the study, the number and selection of subjects, the lack of blind judging and the intentional selection of extreme results, it would be misleading to conduct or report significance tests for this study. Where statistics are shown (e.g., Cronbach's alpha), these indicate criteria used in selecting items for further analysis. As a pilot study, the primary result is intended to be the actual measurement scales rather than statistics relating to their reliability and/or validity. Such estimates would have to be based on a sample separate from that used in derivation of the scales (Cureton 1950).

Emotion Scales

In general, the factor analyses of emotional responses provided encouraging results. By and large, the adjectives which were predicted to describe a particular emotion did load on the same factor. Furthermore, the emotions which had been associated with satisfaction and dissatisfaction also tended to load together (Table 2). As indicated above, the mean score on each scale listed in Table 2 was used in further analyses.

The emotional scales were used to predict whether the subject had been in the satisfied or dissatisfied condition of the experiment (discriminant analysis using PROC GLM from the Statistical Analysis System) and to predict the subject's responses on both the more typical satisfaction scales and the behavioral intent summed scales (regression and correlation in SAS). Two different cluster analysis programs (SAS and BMDP -- K-means) were used to describe subsets of the sample.

No single analysis led to firm conclusions. However, the bulk of the evidence suggested the following conclusions:

1. Izard's abbreviated DES (1977) was more reliable (Table 2) and yielded more consistent results than either Westbrook's version or scales developed within the present study.

2. The emotions which were most characteristic of persons in the "satisfied" condition and were most associated with highly positive satisfaction and behavioral intent responses were: Interest, Joy and Surprise (see Table 2 for specific items used to assess these).

3. The emotions which were most characteristic of persons in the "dissatisfied" condition and were most associated with highly negative satisfaction and behavioral intent responses were: Anger, Disgust and Surprise (Table 2).

These emotions, then, comprise working definitions of either end of the satisfaction continuum. Operational definitions were derived based on mean responses by subjects in each condition. A satisfaction pattern of emotions consists of an average response equal to or greater than 3 (a response of "Moderately") on the items comprising the characteristic emotions. Similarly, dissatisfaction (negative satisfaction) is defined as an average response equal to or greater than 3 on items specific to that pattern of emotions.

TABLE 2
Results of Factor Analysis of Emotion Items

I. Using items from Izard (1977)

Emotion	Coefficient Alpha		Factor Loadings 1	2	3	4	5	6
Joy	.95	Delighted	-.7144					
		Happy	-.7839					
		Joyful	-.6470					
Distress	.77	Downhearted	.7205					
		Sad	.5889					
		Discouraged	.6968					
Disgust	.89	Feeling of Distaste	.8093					
		Disgusted	.9016					
		Feeling of Revulsion	.8139					
Anger	.95	Enraged	.9077					
		Angry	.9221					
		Mad	.9181					
Contempt	.83	Contemptuous	.6923					
		Scornful	.8528					
		Disdainful	.7441					
Interest	.80	Attentive			.8747			
		Concentrating			.7696			
		Alert			.7486			
Fear	.90	Scared		.9404				
		Fearful		.8326				
		Afraid		.9225				
Surprise	.80	Surprised				.8191		
		Amazed				.8274		
		Astonished				.7724		
Shyness	.60	Sheepish					.5707	
		Bashful						.8428
		Shy						.7701
Guilt	.59	Repentant	.4651					
		Guilty					.7851	
		Blameworthy					.8237	

Results from Westbrook's Items (29 items)

Factor 1	Distress, Disgust, Anger, Contempt
Factor 2	Interest, Joy
Factor 3	Surprise
Factor 4	Guilt, Shyness
Factor 5	Fear
Factor 6	Shame

Results from items selected from Izard's Factor Analysis (49 items)

Factor 1	Joy, Distress, Disgust, Anger
Factor 2	Interest
Factor 3	Fear
Factor 4	Surprise
Factor 5	Shyness
Factor 6	(?)*
Factor 7	Guilt

Results from all items (75 items)

Factor 1**	Distress, Disgust, Anger, Contempt
Factor 2**	Joy
Factor 3	Fear
Factor 4	Surprise
Factor 5	Interest
Factor 6	Shyness
Factor 7	Guilt
Factor 8	(?)
Factor 9	(?)
Factor 10	(?)

* (?) indicates that no clear characterization could be made

** Several of the added items (satisfied, dissatisfied, etc.) loaded on these.

When these definitions were compared with the responses of the 94 subjects, 36 met neither cut-off, 7 met both cut-offs, 50 were correctly classified into original condition and 1 subject from the dissatisfied condition was identified as satisfied. All of the "misclassified" subjects were from the dissatisfied condition of the study. Further examination showed that there was a tendency to report "dissatisfying" occasions which were resolved in the subject's favor and resulted in a net satisfaction. This explanation is corroborated by the responses on classical satisfaction scales.

Those subjects who met neither cut-off (14 from the satisfied condition and 22 from dissatisfied) represent a different problem for the measurement technique. Either the definitions that have been constructed are not sufficiently sensitive, or these subjects were not satisfied (or dissatisfied) to the extent that is of interest to CS/D researchers. Although not conclusive, some insight can be gained by comparison against resulting behavioral intentions.

One advantage of the present conceptualization is the integration of specific results (behavioral, motivational) into the overall satisfaction theory. This enables the use of behavioral intent measures as well as satisfaction scales to compare subjects who have been classified by emotional response (Table 3).

Table 4 summarizes the mean responses to various dependent measures by experimental condition or subsequent emotional classification of subjects. There are a couple of points notable from the table: 1) The difference between survey groups (conditions) resulted in significantly different responses across all measures. 2) The Likert type scales used to assess behavioral intention (Table 5) were centered at Neither Agree nor Disagree with a statement indicating what the subject might "normally do" under the circumstances described. The responses to the items comprising summary scales four, five, and six in Table 4 lie fairly close to this response, which had been assigned a value of three, and therefore are not likely to discriminate well between satisfied and dissatisfied (or indifferent) individuals. 3) The emotional classification procedure generated slightly more extreme means than those collected from the entire sample.

While the summed scales are adequate to represent differential responses on the items, it is also of interest to determine which individual items may be useful in further studies. Table 5 summarizes the mean values for each of the dependent measures, by condition and by emotional classification. Those items which exhibit high absolute differences are candidates for inclusion as behavioral intent or behavioral measures in future investigations.

TABLE 3
Behavioral Intent Scales from Factor Analysis (with Loadings)

Item # From Questionnaire	Description	Factors: 1	2	3	4	5	6
1	Satisfied with decision to purchase/use*	.8723					
2	If do over, would feel differently*	-.6750					
3	Choice was wise*	.8467					
4	Feel bad about decision*	-.7634					
5	Did right thing*	.7607					
6	Not happy that I did what I did*	-.6378					
7	Would buy other products of same brand	.5547					
14	Tell a friend good things	.7530					
20	Recommend product to a friend	.7890					
21	Recommend friend avoid product	-.4811		.4720			
22	Publicly endorse product	.5987					
12	Complain to higher authority		.8511				
15	Tell a friend bad things		.7753				
16	Write a manufacturer and complain		.8353				
23	Publicly attack product		.5387				
18	Discard and replace with different brand			.7860			
19	Discard and replace with different product			.7823			
8	Buy other products from same seller				.5869		
13	Compliment higher authority				.7417		
17	Write manufacturer and compliment				.8113		
10	Buy more of product as gift					.8811	
11	Buy similar product as gift					.9058	
9	Buy other products of same type						.7637
24	Fight to remove product from market						-.5973
	Number of items	11	4	3	3	2	2
	Coefficient α (based on student sample)	.95	.83	.85	.72	.91	.47

* From Oliver (1980)

TABLE 4
Mean Responses by Condition and Classification[a]

Measure	Condition		Classification	
	Satisfied (n=42)	Dissatisfied (n=50)	Satisfied (n=29)	Dissatisfied (n=29)
Dependent Measure[a]				
Scale 1[b]	4.35	2.53***	4.39	2.37***
Scale 2	3.70	2.64***	3.73	2.52***
Scale 3	4.27	2.96***	4.34	2.89***
Scale 4	3.31	2.41***	3.46	2.19***
Scale 5	2.96	2.30**	3.19	2.11**
Scale 6	4.12	3.70*	4.14	3.52**
Satisfied-Dissatisfied[c]	6.79	1.77***	6.62	1.69***
Odds	9.29	2.48***	9.24	2.00***
Percent Satisfied	93.10	20.96***	93.45	17.59***
Delighted-Terrible	6.40	2.43***	6.59	2.31***

[a]Some scales have been reversed so that in all cases higher numbers are more positive.

[b]"Scales" are mean responses to behavioral intent scales derived from Table 3

[c]Satisfaction measures are based on those in Table 1.

* p < .05 ** p < .01 *** p < .0001

TABLE 5
Mean Responses to Likert Items by Condition and Classification[a]

Measure	Condition		Classification	
	Satisfied (n=42)	Dissatisfied (n=50)	Satisfied (n=28)	Dissatisfied (n=29)
1. Decision satisfaction[b]	4.67	2.08***	4.72	1.89***
2. Feel Differently	4.26	2.68***	4.21	2.54***
3. Wise choice	4.55	2.32***	4.55	2.07***
4. Feel bad	4.69	3.10***	4.69	3.07***
5. Did right thing	4.50	2.86***	4.52	2.61***
6. Not happy	4.71	3.26***	4.72	3.14***
7. Buy brand	4.10	2.76***	4.21	2.54***
8. Buy from seller	3.93	2.92***	4.00	2.50***
9. Buy product type	3.76	3.60	3.86	3.36
10. More gift	3.00	2.16**	3.24	1.96**
11. Similar gift	2.93	2.44	3.14	2.25*
12. Complain higher	3.93	2.46***	3.93	2.21***
13. Compliment higher	3.31	2.14***	3.48	2.07***
14. Friend/good	4.21	2.14***	4.24	2.04***
15. Friend/bad	3.07	1.96***	3.31	1.86***
16. Complain manufac.	3.71	3.00**	3.52	2.96
17. Compliment manuf.	2.69	2.18*	2.90	2.00**
18. Replace brand	4.24	2.90***	4.31	2.82***
19. Replace product	4.31	3.02***	4.38	2.96***
20. Recommend buy	4.29	2.20***	4.38	2.11***
21. Recommend avoid	3.98	2.32***	4.00	2.25***
22. Endorse	3.55	1.94***	3.66	1.71***
23. Attack	4.10	3.14***	4.17	3.04**
24. Ban	4.48	3.80**	4.41	3.68**

[a]Some scales reversed so that higher value are more positive response.

[b]Measure numbers correspond to item numbers in Table 4

*p < .05 **p < .01 ***p < .0001

DISCUSSION

Overall, behavioral intent, emotion and more standard measures of satisfaction have tended to converge on the distinction between groups. At this juncture, however, it would be inappropriate to attempt to compare the measures competitively. There is not reason to expect divergence in any particular direction, nor was the study designed to provide such tests (a replication should indicate convergence among the measures). Moreover, the scales were developed iteratively using one another as criteria, so no true indicants of reliability are available.

Likert type scales were derived with procedures similar to those usually prescribed. Of special note is that the situations were specifically chosen to create extreme responses on the various measures. Thus, the design of the study selected the extreme groups to use in measure selection, as is generally recommended (Edwards and Kenney 1946; Likert 1932; Shaw and Wright 1967).

The behavioral intent measures may be useful in developing measures that are based on actual behavior. Behavioral measurement of attitude has been attempted, but involves considerable difficulty and cost to execute and, due to the other forces acting on behavior, may be of questionable reliability (Cook and Selltiz 1964; Fishbein 1966). It has been suggested that weighting behaviors may make for more reliable measures (Rosander 1937), but such weighting schemes tend to be unstable (Triandis and Triandis 1965). The best option would seem to be to collect both intent and actual behavior measures and look for convergence. Note that an assumption implicit in the discussion of behavioral measures is that an experimental design will control such factors as cost and effort required for each behavior and, to an extent, expected result. That is, it is expected that perceived costs and benefits of various responses would be equivalent across subjects; only motivation would vary. This would mitigate the arguments presented in the CS/D literature against complaints as a measure of satisfaction.

The study used the critical incident technique that has been used successfully in the job satisfaction literature to revive the memory of particularly salient emotional events. Emotional definitions of the positive and negative endpoints of the satisfaction continuum were derived and shown to be feasible to ascertain with a simple questionnaire. The study also indicated behavioral intent measures which discriminated well between groups of subjects. It remains for future investigations to determine and demonstrate the usefulness of the measurement techniques that have been derived.

REFERENCES

Cardozo, Richard N., "An Experimental Study of Customer Effort, Expectation, and Satisfaction," Journal of Marketing Research, 2(August, 1965), 244-249.

Churchill, Gilbert A. Jr., "A Paradigm for Developing Better Measures of Marketing Constructs," Journal of Marketing Research, 16(February, 1979), 64-73.

Cook, Stuart W. and Claire Selltiz, "A Multiple-Indicator Approach to Attitude Measurement," reprinted in Martin Fishbein (ed.), Readings in Attitude Theory and Measurement (New York: John Wiley & Sons, 1967), 220-235.

Cureton, Edward E., "Validity, Reliability and Baloney," reprinted in William A. Mehrens and Robert L. Ebel (eds.) Principles of Educational and Psychological Measurement (Chicago: Rand McNally & Co., 1967), 271-273.

Day, Ralph L., "The Next Step: Commonly Accepted Constructs for Satisfaction Research," in Ralph L. Day and H. Keith Hunt (eds.) International Fare in Consumer Satisfaction and Complaining Behavior (Bloomington: Indiana University, 1983), 113-117.

_____ and Muzaffer Bour, "Analysis of Average Satisfaction Scores of Individuals Over Product Categories," in Ralph L. Day and H. Keith Hunt (eds.) New Dimensions of Consumer Satisfaction and Complaining Behavior (Bloomington: Indiana University, 1979), 184-189.

Edwards, Allen L. and Kathryn Claire Kenney, "A Comparison of the Thurstone and Likert Techniques of Attitude Scale Construction," reprinted in Martin Fishbein (ed,) Readings in Attitude Theory and Measurement (New York: John Wiley & Sons, 1967), 249-256.

Fishbein, Martin, "The Relationships Between Beliefs, Attitudes and Behaviors," reprinted in David T. Kollat, Roger D. Blackwell and James F. Engel (eds.) Research in Consumer Behavior (New York: Rinehart and Winston, 1970), 216-235.

Izard, Carroll E., Human Emotions (New York: Plenum Press, 1977).

_____, Patterns of Emotions: A New Analysis of Anxiety and Depression (New York: Academic Press, 1972).

Jordan, Kathi A. and James H. Leigh, "Race and Sex Differences of Children in Satisfaction and Other Factors Associated With Video Games," in Thomas Kinnear (ed.) Advances in Consumer Research, Vol. XI (Ann Arbor: Association for Consumer Research, 1984), 94-99.

Likert, Rensis, "The Method of Constructing an Attitude Scale," reprinted in Martin Fishbein (ed.) Readings in Attitude Theory and Measurement (New York: John Wiley & Sons, 1967), 90-95.

Locker, David and David Dunt, "Theoretical and Methodological Issues in Sociological Studies of Consumer Satisfaction With Medical Care," Social Science & Medicine 12(July, 1978), 283-292.

Lundstrom, William J. and Lawrence M. Lamont, "The Development of a Scale to Measure Consumer Discontent," Journal of Marketing Research 13(November, 1976), 373-381.

Maddox, R. Neil, "Two-factor Theory and Consumer Satisfaction: Replication and Extension," Journal of Consumer Research 8(June, 1981), 97-102.

Miller, John A, "Studying Satisfaction, Modifying Models, Eliciting Expectations, Posing Problems, and Making Meaningful Measurements," in H. Keith Hunt (ed.) Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction (Cambridge: Marketing Science Institute, 1977), 72-91.

Mizerski, Richard W., Jon B. Freiden and Robert C. Greene, Jr., "The Effect of the 'Don't Know' Option on TV Ad Claim Recognition Tests," in Richard P. Bagozzi and Alice M. Tybout (eds.) Advances in Consumer Research, Vol. X (Ann Arbor: Association for Consumer Research, 1983), 283-287.

Nunnally, Jum C., Psychometric Theory, Second edition, (New York: McGraw-Hill, 1978).

Oliver, Richard L., "Measurement and Evaluation of Satisfaction Processes in Retail Settings," Journal of Retailing 57(Fall, 1981), 25-48.

_____, "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," Journal of Marketing Research 17(November, 1980), 460-469.

_____ and Robert A. Westbrook, "The Factor Structure of Satisfaction and Related Postpurchase Behavior," in Ralph L. Day and H. Keith Hunt (eds.) New Findings on Consumer Satisfaction and Complaining (Bloomington: Indiana University, 1982), 11-14.

Rosander, A.C., "An Attitude Scale Based Upon Behavior Situations," Journal of Social Psychology 8(1937), 3-15.

Sawyer, Alan G., "Demand Artifacts in Laboratory Experiments in Consumer Research," Journal of Consumer Research 1(March, 1975), 20-30.

Shaw, Marvin E. and Jack M. Wright, Scales for the Measurement of Attitudes (New York: McGraw-Hill, 1967).

Singh, Jagdip and Roy D. Howell, "Consumer Complaining Behavior: A Review and Prospectus," in H. Keith Hunt and Ralph L. Day (eds.) Consumer Satisfaction, Dissatisfaction and Complaining Behavior (Bloomington: Indiana University, 1985), 41-49.

Sobel, Robert S. and Howard McGuire, "Subtractive Models of Satisfaction: Cognitive Processes or Measurement Artifact?" Journal of Psychology 96(July, 1977), 251-258.

Swan, John E. and Linda Jones Combs, "Product Performance and Consumer Satisfaction: A New Concept," Journal of Marketing 40(April, 1976), 25-33.

Triandis, Harry C. and Leigh M. Triandis, "Some Studies of Social Distance," reprinted in Martin Fishbein (ed.) Readings in Attitude Theory and Measurement (New York: John Wiley & Sons, 1965), 199-207.

Wanous, John P. and Edward E. Lawler III, "Measurement and Meaning of Job Satisfaction," Journal of Applied Psychology 56(Number 2, 1972), 95-105.

Webb, Eugene J., Donald T. Campbell, Richard D. Schwartz and Lee Sechrest, Unobtrusive Measures: Nonreactive Research in the Social Sciences (Chicago: Rand McNally College Publishing Co., 1946).

Westbrook, Robert A., "Consumer Satisfaction and the Phenomenology of Emotions During Automobile Ownership Experiences," in Ralph L. Day and H. Keith Hunt (eds.) International Fare in Consumer Satisfaction and Complaining Behavior (Bloomington: Indiana University, 1983), 2-9.

_____ and Richard L. Oliver, "Developing Better Measures on Consumer Satisfaction: Some Preliminary Results," in Kent Monroe (ed.) Advances in Consumer Research, Vol. VIII (Ann Arbor: Association for Consumer Research, 1981), 94-99.