

CONFOUNDS IN THE MEASUREMENT OF PREDICTIVE EXPECTATIONS

Richard A. Spreng, Michigan State University
Robert D. Mackoy, Butler University
Cornelia Droge, Michigan State University

ABSTRACT

Given the importance of predictive expectations in consumer satisfaction models, confounds in the measurement of expectations could result in misspecified models. Results of two empirical studies indicate that consumers interpret the word "expect" in numerous ways. A large minority of consumers interpret "expect" to mean "desire." The magnitude of the resulting confounding effect is illustrated by comparing results using a measure of expectations alone with results obtained when using a measure of expectations together with a measure of desires in a side-by-side format.

INTRODUCTION

Consumer satisfaction continues to be a critical area of academic research and managerial interest. While the disconfirmation of expectations model has dominated research, new models and approaches have been suggested (e.g., Woodruff, Cadotte, and Jenkins 1983; Spreng, MacKenzie, and Olshavsky 1996; Oliver and Swan 1989). Despite the great amount of research that has tested the disconfirmation of expectations model, disagreement remains concerning definitions and measurement of some key concepts in consumer satisfaction research (Yi 1990).

In particular, the "expectations" concept has been defined and operationalized in a variety of ways, and we believe that there are two problems with the use of "expectations" in past research. First, there is a disagreement regarding the conceptual definition of expectations. In some cases expectations are viewed as predictions of future product performance, often conceptualized as a likelihood of occurrence (e.g., Bearden and Teel 1983; Olson and Dover 1979; Westbrook 1987; Westbrook and Reilly 1983). Others have argued that expectations involve both an estimate of the likelihood of an event, and an evaluation of how good/bad the event is (e.g., Churchill and Surprenant 1982; Oliver 1980; Tse and Wilton

1988).

An example of this latter perspective is provided by Oliver (1981, p. 33):

"Expectations have two components: a probability of occurrence (e.g., the likelihood that a clerk will be available to wait on customers) and an evaluation of the occurrence (e.g., the degree to which the clerk's attention is desirable or undesirable, good or bad, etc.). Both are necessary because it is not clear at all that some attributes (clerks, in our example) are *desired* by all shoppers." [emphasis added]

As Oliver's discussion makes clear, this conceptualization confounds a person's judgment of the desirability of something with his/her expectation of the likelihood of its occurrence. Additional research highlights potential confounds other than "desires." For example, Zeithaml, Berry and Parasuraman (1993) hypothesize that a third type of expectation is relevant in service settings: the minimally adequate level of service. Might not some respondents in some contexts reasonably interpret "expectations" in this manner, too? In fact, this ambiguity can be found in dictionary definitions of "expect," in that both an "anticipate" and a "desire" definition are given, as well as normative definitions.

Different conceptualizations of "expectations" is a serious problem given its role in models of satisfaction and service quality. For example, it has been demonstrated that desires are distinct from predictive expectations and influence perceptions of quality and customer satisfaction differently. Spreng and Olshavsky (1993) provide both conceptual and empirical evidence that predictive expectations and desires have distinctly different roles in satisfaction formation, while Zeithaml, Berry and Parasuraman (1993) argue that multiple "types" of expectations, including predictive and desired, are relevant in service contexts. Boulding et al (1993) differentiate between "will" expectations and "should" expectations, where the former is predictive

expectations and the latter is a type of normative expectations, described as constrained ideal or desired expectations. They find that each affects perceptions of quality differently: "will" expectations are positively related to perceptions of quality while "should" expectations are inversely related to perceptions of quality. Not only do expectations and desires have differing effects within satisfaction and service quality modeling, there is some evidence that their use as comparison standards also produces differential effects. Spreng and Mackoy (1996) found that while expectations disconfirmation had a significant effect on overall satisfaction, desires congruency influenced both satisfaction and perceived service quality.

The second problem is related to this ambiguity in that it is probable that consumers will also be confused in answering questions about their "expectations." Some consumers may adopt a "predictive expectations" interpretation of the question, some may use a "desires" interpretation, while still others may use a "normative" interpretation. Thus, when researchers ask consumers about their "expectations" regarding a product or service, we believe consumers will use multiple interpretations. If this is true, a great deal of research investigating the role of expectations and disconfirmation of expectations as an antecedent of satisfaction would be called into question. This type of confound may help explain why attempts to measure the effects of predictive expectations on satisfaction formation have yielded inconsistent results. Some researchers have found that expectations and/or disconfirmed expectations are significant antecedents of satisfaction (Bearden and Teel 1983; Churchill and Surprenant 1982 [plant model]; Tse and Wilton 1988) while others have not (Spreng and Olshavsky 1993; Churchill and Surprenant 1982 [video recorder model]; Barbeau 1985). Thus, it is possible that at least some of the inconsistency may be due to respondent interpretation of the term "expectation" or "expect": if some respondents interpret "expect" to mean "predict" or "anticipate," while others interpret it to mean "desire," it seems reasonable that results could be confounded.

Therefore, the purpose of this research effort is to 1) determine the degree to which consumers use alternative definitions of expectations, 2)

determine the extent to which any confusion may affect the measurement of expectations, and 3) investigate one alternative method of minimizing such confusion if it exists.

If predictive expectations are confounded with desires, actual relationships between expectations and post-consumption variables will be confounded. Such a finding would call into question much of customer satisfaction modeling research, which has relied extensively on the disconfirmation of expectations paradigm and which has not typically included measures of desires, as well as much service quality literature, which has not consistently included measures of predictive expectations. If expectations and desires each affect satisfaction independently, and expectations are confounded with desires, then the problem will be especially serious for studies in which only one or the other is measured. Only a handful of studies have included measures of both predictive expectations and desires (Westbrook and Reilly 1983; Barbeau 1985; Tse and Wilton 1988; Spreng and Olshavsky 1993; Spreng, MacKenzie, and Olshavsky 1996; Spreng and Mackoy 1996); these studies found that expectations and desires had different effects on satisfaction.

The exploratory research effort reported here consisted of two studies. In study 1, we attempted to determine explicitly which definition of expectations was used by people who were asked to indicate their expectations in common consumption contexts. In study 2, we focused on the degree to which measurement of predictive expectations and desires may be confounded.

STUDY 1

Method

Four hundred thirty three students in an introduction to marketing class were asked to complete a brief (less than 5 minute) survey. Students were asked to imagine a common consumption situation such as going to McDonald's for lunch, purchasing an airline ticket, buying a Coca-Cola, purchasing a Ford Escort, etc.; each student was presented with only one situation. Students were asked to indicate on a Likert scale the degree to which they expected

the target product/service to possess specific features. For example, those presented with the McDonald's scenario were asked to indicate how strongly they agreed or disagreed that McDonald's would "... be clean," "... have fast service," "... have a friendly staff," "... be inexpensive," etc. Responses were recorded on seven point scales.

Once this simple task was completed, students were asked to complete three questions on the back of the questionnaire. Instructions prior to the three questions explained that multiple definitions for the word "expect" exist, and that "people often interpret the word in different ways." Students were then asked which of four possible definitions of expect was most similar to the definition they personally used to respond to the earlier scenario.

The first question read:

Check the one interpretation of expectations which is closest to the interpretation you actually used to answer the questions above.

The characteristics I feel that I must receive.

The characteristics I want to receive.

The characteristics I feel would be minimally adequate.

The characteristics I believe I will actually receive.

Other: The characteristics I _____.
(Use your own words to explain your interpretation.)

We recognize that subjects may use different definitions of expectation in different situations. Therefore the other two questions asked students which definition of expectations was most applicable to them personally when confronted with a familiar product and which definition was most applicable when confronted with an unfamiliar product.

Results

All 433 students returned completed questionnaires. Simple frequencies were tabulated for each of the possible definitions of expectations. Responses were nearly equally divided across all

four possible responses:

....feel that I must receive 24%

....want to receive 24%

....feel would be minimally adequate 23%

....believe I will actually receive 24%

....other 4%

Similar patterns appear for the two remaining questions, with no category being selected by more than 28% or fewer than 21% of the respondents. Further, the majority of respondents indicated that they use different interpretations of the term "expect" in different situations: only 15% reported using the same interpretation of "expectations" for all three questions. This is a significantly higher proportion ($p < .01$) than the 2% expected by chance, but still extremely low. A three-way cross-tabulation analysis failed to reveal any meaningful pattern in the data. Thus, not only does interpretation of "expectations" differ between subjects but also within the same subject.

STUDY 2

Methods

Consequences of confounding predictive expectations and desires may not always be apparent in the results of empirical investigations, especially those of field studies. Product and service providers expend considerable effort trying to produce products/services which match consumer desires, and then try to raise customer expectations to these levels. In many product/service contexts, therefore, predictive expectations and desires are very similar, and it is unlikely that measurement-related confounds in these contexts would be evident. Thus, a test of the existence (and strength) of the confound should be conducted under conditions in which desires and predictions are likely to be similar as well as dissimilar.

Data were collected from undergraduate business students enrolled at a large midwestern university. Participation was voluntary and no student declined to participate. The study focused on the undergraduate student advising center, a service with which most students were familiar. A

brief questionnaire was administered in a classroom setting and consisted of three parts. Part 1 contained basic classification questions, such as year in school, age, and gender. Part 2 was an expectations manipulation, designed to engender low versus high service expectations regarding the advising center. Part 3 contained questions that either measured expectations alone, or measured expectations and desires in a side-by-side format. A total of 174 students completed a questionnaire in the 2 (expectations) X 2 (measurement format) design.

Expectations were manipulated in Part 2 of the questionnaire by exposing subjects to one of two ads purportedly from the advising center. The ads represented realistic information about the advising center, with one ad intended to lower expectations, while the other was intended to raise expectations.

The two different expectations measures are referred to as 1) "traditional" expectations measure, and 2) "juxtaposed" expectations measure. The "traditional" measure included the word "expect" and "expectations" several times in the instructions. The "juxtaposed" measure required subjects to indicate their desired level of service followed by their expected level of service for each attribute. All scales were 7-point "strongly disagree" (1) to "strongly agree" (7). Specific attributes are listed in Table 1.

Table 1
Description of the Attributes

Attribute Number	Description
1	convenience in making an appointment
2	friendliness of staff
3	advisor listened to questions
4	advisor provided accurate information
5	knowledge of advisor
6	advice was consistent
7	advisor helped in long range planning
8	advisor helped in choosing right courses for career
9	advisor was interested in my personal life
10	advising offices looked professional

The following hypotheses were tested in study 2.

H1: Average expectations ratings under the

positive manipulation will be higher than the average expectations ratings under the negative manipulation.

H2: Average desires ratings under the positive expectations manipulation will be equal to the average desires ratings under the negative expectations manipulation.

H3: Expectations, when measured alone (traditional), will yield average ratings which are higher than those yielded when expectations are measured with desires (juxtaposed).

Hypothesis 1 and Hypothesis 2 are straightforward. Hypothesis 3 is the focus of this analysis. The rationale for Hypothesis 3 is that the traditional measure of expectations will be confounded as some subjects will interpret expectations in terms of their desired level of service, while others will interpret expectations in terms of the level of service they actually expect to receive. In other words, traditional measures of expectations ought to fall between measures of desires and the juxtaposed measures of expectations. Support for the hypothesis would be consistent with our contention that traditional measures of expectations are actually "weighted averages" of various interpretations of expectations, and not merely averages of predictive expectations across respondents (which is what researchers often think they are measuring).

Results

Hypotheses 1, 2, and 3 were supported based on t-test difference of means analysis.

To test hypothesis 1, twenty difference of means t-tests were conducted, ten comparing the traditional expectations measures in the positive versus negative manipulation condition (for each attribute), and ten comparing the juxtaposed expectations measures in the positive versus negative manipulation conditions (for each attribute). For the tests using traditional expectations measures, expectations in the positive condition were significantly higher than expectations in the negative condition for all ten attributes ($p < .01$, one-tailed tests). For the tests

Figure 1
Traditional Expectations, Juxtaposed Expectations, and Desires
Positive Manipulation

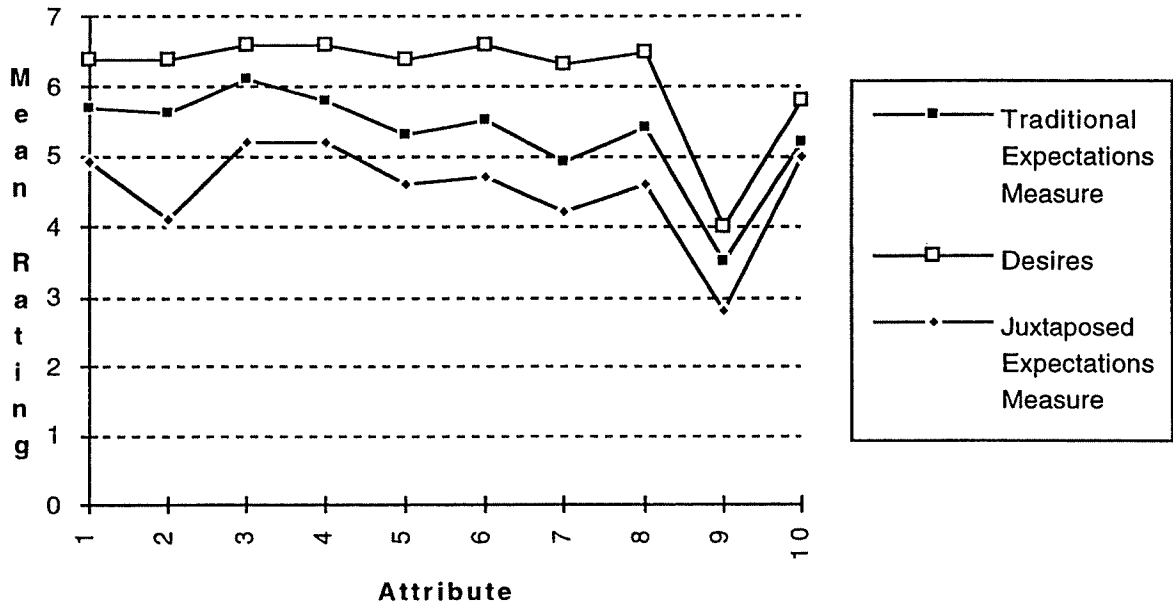
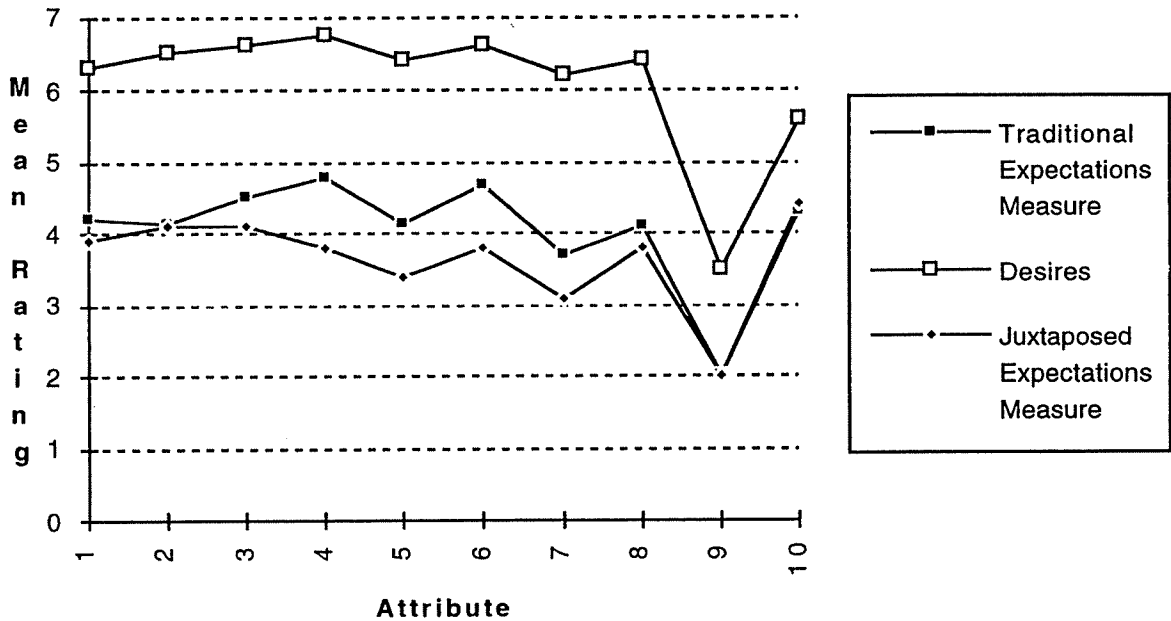


Figure 2
Traditional Expectations, Juxtaposed Expectations, and Desires
Negative Manipulation



using juxtaposed expectations measures, expectations in the positive condition were significantly higher than expectations in the negative condition for nine of the ten attributes (p

$< .025$, one-tailed tests); for attribute #2, the juxtaposed measures of the positive versus negative manipulation condition were equal. Thus, we found strong support for hypothesis 1.

To test hypothesis 2, ten difference of means t-tests were conducted comparing desires measures for each attribute under the positive versus negative expectations manipulation. No difference was statistically significant (lowest p-value = .191, two-tailed test). Therefore, hypothesis 2 was also strongly supported.

To test hypothesis 3, the mean expectation levels yielded by the two different measures of expectations (traditional versus juxtaposed) were compared on an attribute-by-attribute basis using the t-test difference of means. Figure 1 illustrates the traditional expectations, juxtaposed expectations and desires mean ratings for the positive manipulation, while Figure 2 illustrates means of the same ratings for the negative manipulation. In the positive manipulation condition, the pattern of expectations measures is as hypothesized, that is, traditional measures of expectations lie between juxtaposed expectations and desires for every attribute. The mean traditional measure of expectations is significantly higher ($p < .05$) than the mean juxtaposed measure of expectations for every attribute except #4 and #10. Likewise, in the negative manipulation condition, the pattern of expectations measures is as hypothesized for all attributes except #2, #9, and #10. However, the mean traditional measure is significantly higher ($p < .05$) than the mean juxtaposed measure for only attributes #4, #5, #6, and #7. Note that the mean juxtaposed measure was not significantly higher than the mean traditional measure for any attribute in either condition.

The results offer strong evidence that people do use different interpretations of expectations. The traditional measure of expectations does appear to be confounded: its value across multiple attributes under both conditions is consistent with the proposition that some people used a "desires" interpretation while others used a "predictive" interpretation of expectations.

DISCUSSION

The extent to which respondent confusion between predictive expectations and desires has affected previous research is difficult to assess. Exact question wording is generally not reported so it is impossible to determine the extent to which

the term "expect" or "expectation" is actually used in questions designed to measure predictive expectations. In addition, it may be true that different contexts may have different effects on the amount of any confusion. For example, the context of durable goods may elicit a higher (or lower) proportion of respondents to interpret expectations as desires relative to consumer goods. Likewise, services which are familiar may elicit a higher (or lower) proportion of respondents to interpret expectations as desires relative to services which are unfamiliar.

One implication, not tested in this study, is that differing interpretations of "expect" may have an impact on measures of subjective disconfirmation. Disconfirmation is usually measured on a scale ranging from "much better than expected" to "much worse than expected." Thus, even when predictive expectations are accurately measured, measures of disconfirmation may be subject to the same types of confounds as discussed in this paper. Given the wide use of subjective disconfirmation in satisfaction modeling, a systematic confound associated with this construct could be an additional serious problem.

CONCLUSIONS

The terms "expect" or "expectations" appear to be ambiguous. At the very least, the terms do not discriminate between the concepts of "predictive expectation" and "desires." As one might expect, the problem appears to have more severe consequences when predictive expectations and desires are likely to be far apart.

One clear implication for both researchers and managers is that the term "expectation" (or "expect") should be avoided if possible in questionnaires. If the researcher or manager wants to measure predictive expectations, "anticipate actually receiving" could be used. Given that both desires and predictive expectations may be relevant in service quality or satisfaction formation, measuring both constructs in a juxtaposed format appears to be acceptable as this method appears to discriminate between the two constructs.

REFERENCES

- Barbeau, J. Bradley (1985), "Predictive and Normative

- Expectations in Consumer Satisfaction: A Utilization of Adaption and Comparison Levels in a Unified Framework," *Consumer Satisfaction, Dissatisfaction, and Complaining Behavior Conference Proceedings*, H. Keith Hunt, (ed.), Provo, UT: 27-32.
- Bearden, William O. and Jesse E. Teel (1983), "An Investigation of Personal Influences on Consumer Complaining," *Journal of Marketing Research*, 20, (February): 21-28.
- Boulding, William, Ajay Kalra, Richard Staelin, and Valarie A. Zeithaml (1993), "A Dynamic Process Model of Service Quality: From Expectations to Behavioral Intentions," *Journal of Marketing Research*, 30, (February), 7-27.
- Churchill, Gilbert A., Jr. and Carol Surprenant (1982) "An Investigation into the Determinants of Consumer Satisfaction," *Journal of Marketing Research*, 19, (November), 491-504.
- 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. (1981), "Measurement and Evaluation of Satisfaction Processes in Retail Settings," *Journal of Retailing*, 57, (Fall), 25-48.
- 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.
- Olson, Jerry C. and Philip A. Dover (1979), "Disconfirmation of Consumer Expectations Through Product Trial," *Journal of Applied Psychology*, 64:2, 179-189.
- Spreng, Richard A., Scott B. MacKenzie and Richard W. Olshavsky (1996), "A Reexamination of the Determinants of Consumer Satisfaction," *Journal of Marketing*, 60, (July), 15-32.
- Spreng, Richard A. and Richard W. Olshavsky (1993), "A Desires-Congruency Model of Consumer Satisfaction," *Journal of the Academy of Marketing Science*, 21, (Summer), 169-177.
- Spreng, Richard A. and Robert D. Mackoy (1996), "An Empirical Examination of a Model of Perceived Service Quality and Satisfaction," *Journal of Retailing*, 72, (Summer), 201-214.
- 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. (1987), "Product/Consumption-Based Affective Responses and Postpurchase Responses," *Journal of Marketing Research*, 24, (August), 258-270.
- Westbrook, Robert A. and Michael D. Reilly (1983), "Value-Percept Disparity: An Alternative to the Disconfirmation of Expectations Theory of Consumer Satisfaction," *Advances of Consumer Research*, 10, Richard P. Bagozzi and Alice M. Tybout, (eds.), Ann Arbor, MI: Association for Consumer Research, 256-261.
- 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.
- Yi, Youjae (1990), "A Critical Review of Consumer Satisfaction," *Review of Marketing 1990*, Valarie A. Zeithaml, (ed.) Chicago, IL: American Marketing Association.
- Zeithaml, Valarie A., Leonard L. Berry and A. Parasuraman (1993), "The Nature and Determinants of Customer Expectations of Service," *Journal of the Academy of Marketing Science*, 21, (Winter), 1-12.

Send correspondence regarding this article to:

Robert D. Mackoy
College of Business Administration
Butler University
4600 Sunset Avenue
Indianapolis, IN 46208 USA

REFERENCE AND MEMORY IN CUSTOMER SATISFACTION STUDIES

James H. Drew, GTE Laboratories Inc.

ABSTRACT

Revealed customer opinion is not necessarily static over time for a variety of reasons, many of which are functions of consumer memory and not directly related to objective levels of service provision. In the absence of empirical data, corporate folklore has developed several divergent views of the effect of customer memory and past experience on current satisfaction levels. Through a general statistical model of service quality evaluation and its change dynamics, objective mediation among these competing hypotheses is possible. This paper presents two longitudinal studies of divergent customer bases whose data modeling sheds light on several aspects of customer memory as it relates to service quality evaluation and modification.

INTRODUCTION

Customer satisfaction researchers have developed extensive theories for the nature and formation of the satisfaction affect at a single point in time. Most companies, though, record and track quality and satisfaction statistics over time, often through survey designs which gather periodic data from the same customers. It is then possible to examine and model how satisfaction and quality judgments change over time and in relation to which stimuli. In particular, since both sorts of affect depend on reference standards, it is important to also model the nature, time lag and dynamics of such references.

Opinion Updating Models

We will examine some of the models of opinion dynamics which have been developed in the marketing, cognitive psychology and survey research literature. Insofar as many of these models were constructed using laboratory or non-experiential data, the results reported here afford tests of some well-known models and opportunities for their refinement.

Models from Marketing Science. There are

several models in the marketing literature which explain changes in customer satisfaction through successive updates in expectation levels. Drawing on the adaptation level theory of Helson (1964), LaBarbera and Mazursky (1983) tested a model in which prior satisfaction standards are adapted to current service experience, thus creating an updated standard. Boulding et al, (1993) have developed a fuller model of the interplay between customers' expectations, perceptions and evaluations, and how these constructs are updated over successive service experiences.

In their model, "would" (i.e. predictive) expectations, and "should" (i.e. normative) expectations are updated over successive service experiences, and evaluated along with service attributes to produce a customer satisfaction level via the comparison of expectations and performance. Delivered service at time t , for example, is positively related to the "would"-expectation at time $t+1$, which in turn is positively related to delivered service at time $t+1$. This model thus explains correlations in customer perceptions between times t and $t+1$ through delivered service at time t .

Note that although both the Boulding model and the adaptive expectations model will behave similarly in the sense that a prior positive experience will have a positive effect on the current rating, and a negative one will depress it, the former model links consecutive aggregate ratings only through delivered service attributes (which drive the updating of expectations), while the latter allows for a direct link.

Models from Cognitive Psychology. We have outlined two adaptation models already in use in the marketing tradition. An even more explicit model for belief updating was developed by Hogarth and Einhorn (1992), which draws on research in cognitive psychology. This model was constructed to test the classic primacy model, where the current rating is completely dependent on any initial rating (e.g. Nisbett and Ross, 1980), and the classic recency model, where the current rating completely supersedes any prior ratings (e.g. Davis, 1984).

They postulate an anchor-and-adjustment model, wherein attitudes are updated with each discrete new experience X_t , and current attitude A_t is a weighted average of the prior attitude and the difference between the current experience and one of several reference points:

$$A_t = A_{t-1} + w_t (X_t - R).$$

In this formulation, R can be a constant, rather like a "should" expectation, in which case the ordering of the experiences is irrelevant and there are no primacy or recency effects; or R can depend on the prior attitude A_{t-1} itself, in which case the updating is like a moving average and the order of experiences is quite important. Further, the weight w_t of the second, "disconfirmation" term in parentheses is related to the size and parity of the disconfirmation $EX_t - R$: if the experience is more favorable than the reference, then the weight is inversely proportional to the anchoring attitude A_{t-1} , and if the experience is less favorable then it is directly proportional to A_{t-1} . This is a contrast effect, and specifies the notion that adjustments are larger in the direction toward $EX_t - R = 0$.

Models from Survey Research. Two contrasting forms for attitude responses have been mentioned in the survey literature. Waterton and

Lievesley (1989) mention two such models. One is that repeated interviewing "freezes" attitudes, or that repeated interviewing speeds up the process of forming a stable attitude, so that attitudes are consistent over observations. The second postulates raised consciousness through repeated interviewing, so that adjustments in attitude, particularly from the first to the second interview period are common. In their analysis of the British Social Attitudes Survey, Waterton and Lievesley find no strong evidence of either effect.

Other models from the survey literature focus directly on the role of the instrument. Simmons et al. (1993) discuss the role of the questionnaire in creating attitudes not existing prior to the subject's response. They also postulate conditions under which responses to one item may affect responses to a subsequent item. Schul and Schiff (1993) also address the issue of inter-item effects, and develop an argument for the longevity of the memory of inadequate service encounters. This is a restating of the folklore that customers remember service experiences for long time periods, and that a service problem is remembered, and has an effect on overall opinion for a longer time than service without incident.

These models, and some of their salient or distinguishing features, are summarized in the following table:

Model	Source	Important Features
Adaptive Expectations	Helson (1964)	Direct linkage of aggregate rating
"Would"/"Should" Expectation Updating	Boulding et al. (1993)	Rating linkage through expectations, service attributes, <u>not</u> aggregate ratings
Primacy	Nisbett and Ross (1980)	First rating determines subsequent ratings
Recency	Davis (1984)	Prior ratings completely overcome by current rating
Anchor/Adjustment	Hogarth-Einhorn (1992)	Adjustment amount from previous reference point depends on reference level
Memory Freezing	Waterton and Lievesley (1989)	No attitude change from early ratings
Respondent Sophistication	Waterton and Lievesley (1989)	Largest change in attitude between first and second survey wave

Data From Study I

The data that we analyze in this study were edited from a large set of business customer opinion survey interviews conducted from 1989-1992. These surveys make annual contact with identified telecommunications decision makers in medium and large businesses in telephone company franchise areas. Substantial editing was necessary to delete businesses whose interviews were not matched during this period. Further, provision had to be made for separate interviews for a given business which happened to be conducted in the same calendar year. (This can occur through errors in the survey frame or in interviewing schedules.) We assume such deletions are not related to survey responses or their interactions. Previous studies of cross-sectional data have found, in fact, that for these types of customer opinion surveys, there is no measurable effects due to nonresponse; see Drew (1990). The combined dataset contained approximately 3000 customers who furnished two or more interviews during this time period.

Anchor-Adjustment Models. A popular model for attitude formation, mentioned by Kahneman and Tversky (1973), makes the current attitude an adjustment from some form of reference value. Updates in attitude are due to changes in current perceived performance and an updating of the reference value from which current performance generates an adjustment. The updating of the reference level has the form:

$$\begin{aligned} \text{REF}_{t+1} &= \text{REF}_t + (1 - \rho) (x_t - \text{REF}_t) \\ &= \rho \text{REF}_t + (1 - \rho) x_t \end{aligned}$$

where REF_t is a reference value perhaps depending on previous experience, x_t is a measure of current (i.e. time t) perception, and ρ is a weight measuring the relative contribution of x_t and REF_t . Note that $\rho = 1.0$ yields the recency model, while $\rho = 0.0$, with an initial reference value set equal to the first rating yields the primacy model, with the added condition that no other current attributes enter the model as explanatory variables. Hence, these two models can be evaluated in the context of the anchor-adjustment model.

As mentioned earlier, in the Hogarth-Einhorn

model, the weight ρ is taken to be directly proportional to the reference value when the current rating exceeds the reference value, and inversely proportional to the reference value when the current rating is smaller.

An extended model can also incorporate the features of the memory freezing model. By postulating that the reference value will not be updated unless the current perception is much more favorable or much less favorable, one can simultaneously allow freezing, as well as the Hogarth-Einhorn form of reference adjustment. This is accomplished by allowing these two different models to operate in mutually exclusive rating response intervals. As noted above, the Boulding model can be tested by allowing the inclusion of attributes from prior time periods.

In our data, the rating at hand is QREP, the customer's evaluation of overall repair quality, and its significant attributes were found (from earlier studies) to be RFX and RDEAD (Fixing the Problem the First Time, and Meeting Deadlines, respectively.)

Consider the specification below, taken from Drew and Bolton (1996). The basic form of this model is anchor-adjustment as postulated by Hogarth-Einhorn, with the anchor for time t represented by REF_t and the adjustment a function of repair attributes for the current (time t) period. The anchor REF_t consists of a weighted average of the preceding general rating QREP_{t-1} and the preceding anchor, and the attribute ratings of the preceding period, as the Boulding model and a test of memory sophistication require. The weighting function of the preceding anchor and general rating take on a slightly more general form than suggested by Hogarth-Einhorn: first, non-zero weighting takes place only when the preceding rating differs from the preceding anchor by more than k units; and second, the weight is based on a power of the preceding reference. Hogarth-Einhorn specify that $k=0$ and $\alpha_+ = -\alpha_- = 1.0$, and a natural alternative is $\alpha_+ = \alpha_- = 0.0$.

$$\begin{aligned}
 \text{QREP}_t = & \begin{cases} \text{REF}_t + \beta_1 \text{RFX}_t + \beta_2 \text{RDEAD}_{t1} + \epsilon_{it} & \text{if } \text{QREP}_{t-1} - \text{REF}_{t-1} \geq -k \\ \text{REF}_{t-1} + \beta_1 \text{RFX}_t + \beta_2 \text{RDEAD}_{t-1} + \epsilon_{it} & \text{if } |\text{QREP}_{t-1} - \text{REF}_{t-1}| < k \\ \text{REF}_t + \beta_1 \text{RFX}_t + \beta_2 \text{RDEAD}_{t-1} + \epsilon_{it} & \text{if } \text{QREP}_{t-1} - \text{REF}_{t-1} < k \end{cases}
 \end{aligned}$$

where

$$\text{REF}_t = \rho_- \text{REF}_{t-1} + (1-\rho_-) \text{QREP}_{t-1}, \text{ if } \text{QREP}_{t-1} - \text{REF}_{t-1} \geq -k,$$

$$\text{REF}_t = \rho_+ \text{REF}_{t-1} + (1-\rho_+) \text{QREP}_{t-1}, \text{ if } \text{QREP}_{t-1} - \text{REF}_{t-1} < k, \text{ and}$$

$$\rho_- = \beta_- \text{REF}_{t-1}^{\alpha_-} \text{ and } \rho_+ = \beta_+ \text{REF}_{t-1}^{\alpha_+}$$

Since the Hogarth-Einhorn model includes recency and primacy as special cases, and the inclusion of QREP_{t-1} in the anchor subsumes adaptive expectations, our model thus includes the characteristics of each of the rating change models given in the table above, except perhaps for the freezing and sophistication of memory, which do not seem viable in this context. Note too that by allowing the form of the current reference to be a function of the prior survey variable, we capture the notion that responses are potentially artifacts of the item form., that is, that QREP_t should look much like QREP_{t-1}.

Since the join points of this segmented regression are unknown, this model must be fit by a nonlinear routine. The complexity of this model, along with the response patterns of the 162 customers sampled, did not allow the unrestricted fitting of α₋ and α₊, so that only the two alternatives α₊ = -α₋ = 1.0 and α₊ = α₋ = 0.0 were fit and compared. The resulting loss functions were nearly identical, and there is no evidence to reject the Hogarth-Einhorn form, but for our purposes it is more interpretable to allow α₊ = α₋ = 0.0. The period t=3 is tested, and to test the constancy of the anchor the initial anchor REF₂ is given the linear form

$$\mu_0 + \beta_0 \text{QREP}_1.$$

The model was simplified in other ways. Fitting a non-zero value for k invariably resulted in estimates for μ₀ + β₀ QREP₁ and k which led to

sparse or empty model segments. Interpreted as an indication that these data do not support three segments, k was fixed at zero. As noted below, the current attributes were also ignored in the results reported here. Further, a preliminary model gave nearly identical estimates for β₋ (=ρ₋) and β₊ (=ρ₊) and the common value is denoted by ρ.

A simplex routine (the O'Neill, 1971 algorithm as implemented in SYSTAT) yields the following least squares estimates when t = 3, and α₊ = α₋ = 0, which were confirmed using several different starting values for the algorithm, and where the estimated asymptotic standard errors are given in parentheses:

Parameter	Estimate (a.s.e.)
μ ₀	0.880 (0.236)
β ₀	0.150 (0.060)
ρ	0.968 (0.061)
β ₁	0.271 (0.061)
β ₂	0.323 (0.065)

All coefficients are significantly different from zero at the 0.001 level. It is possible, then, to

comment on aspects of customer memory that have been previously discussed. Our estimates imply that:

! QREP_t does not depend on its value from the preceding time period nearly as much as it depends on current period attributes.

! There is no evidence that k is significantly larger than zero, so the updating mechanism is uniform across waves.

! The single anchor update weight is close to 1.0. Therefore, updating of the anchors heavily favors the initial anchor. Further, the initial anchor is only slightly dependent on historic perceived service, and thus not far from being a constant.

! The overall repair rating QREP is unlikely to be a survey artifact, in the sense that the previous period QREP variable and the original anchoring value QREP₁ have little influence on the current period rating.

Finally, it is important to note some indications from unsuccessful estimation runs. Models with both the updating mechanism and the current period attributes RFX₃ and RDEAD₃ generally either did not converge to estimates with reasonably small asymptotic standard errors, converged to estimates with k very large or very small, converged to estimates with the attribute coefficients indistinguishable from zero, or did not converge at all. Presumably, this shows that updated past evaluations carry about the same information as current attribute ratings. This, of course, has important implications for the use of the panel designs that are necessary to acquire time series data from customers.

Data from Study II

The preceding results should perhaps be viewed cautiously, on the grounds that survey responses were collected no more frequently than at annual intervals, that POTS (Plain Old Telephone Service) evaluation is not cognitively stimulating, or that the memory lag and updating structure is specific to business customers only.

The following study is based on three waves of telephone surveys with cellular telephone customers surveyed during the time periods January-March 1992, August-September 1992 and April-May 1993. Unlike the business customers of our first study, these customers were mostly personal and small-business users, and few made telecommunications decisions for anyone other than themselves.

A small number of customers were dropped from the analysis because they chose to terminate service. Most such termination was not related to issues of cellular quality, and a logistic regression revealed that termination was not statistically related to prior quality ratings. The number of customers interviewed for all three waves of the study was 245.

The questionnaires from this survey program each contained items associated with service quality and its attributes; the following items ask for the overall satisfaction (on a 1-5 scale) with:

OVQ_i: the overall services you received
 MAKE_i: the calls you make and receive
 FIRST_i: calls going through the first time
 PRICE_i: the prices charged,
 CLEAR_i: the clarity of the calls, and
 CALLCS_i: indicates whether a call was made to Customer Service.

The subscript refers to the survey wave ($i=1,2,3$) in which the data were furnished.

Consider first this simplification of the model introduced above:

$$OVQ_3 = \begin{cases} (1-\rho_-) OVQ_2 + \rho_- (\mu_0 + \beta_0 OVQ_1) , & \text{if } OVQ_2 < \mu_0 + \beta_0 OVQ_1 - k \\ (1-\rho_+) OVQ_2 + \rho_+ (\mu_0 + \beta_0 OVQ_1) , & \text{if } OVQ_2 > \mu_0 + \beta_0 OVQ_1 + k \end{cases}$$

In a preliminary analysis, it was found that k was effectively zero, and there was no significant difference between the reference weights ρ_- and ρ_+ in the two model segments. In the full model below, then, we set $k=0$ and $\rho_- = \rho_+ = \rho$.

With this simplification, it is feasible to fit a fuller model with the current-period quality attributes listed above:

$$OVQ_3 = \beta_{Make} MAKE_3 + \beta_{First} FIRST_3 + \beta_{Clear} CLEAR_3 + \beta_{Price} PRICE_3 + \beta_{CallCS} CALLCS_3 + \rho OVQ_2 + (1-\rho) (\mu_0 + \beta_0 OVQ_1)$$

The estimated parameters are:

Parameter	Estimate (Standard Error)
β_{Make}	0.447 (0.066)
β_{First}	0.123 (0.061)
β_{Clear}	0.026 (0.050)
β_{Price}	0.065 (0.039)
β_{CallCS}	-0.073 (0.097)
ρ	0.041 (0.045)
μ_0	1.181 (0.296)
β_0	0.071 (0.056)

Note again that the once-lagged quality rating OVQ_2 has only a small effect on the reference value, which is almost a constant. The reference value itself is small in relation to the contribution of the current-period quality attributes $MAKE_3$ and $FIRST_3$. This is a finding similar to that of Study I.

Further information on reference points is afforded by analysis of the variable $EXPSE_{i-1}$ ($i=2,3$), which compares overall service perception with expectation. This is effectively the

expectancy-disconfirmation measure that Oliver (1980) and others identify as a distinguishing ingredient in the customer satisfaction affect.

SUMMARY

We have considered the modeling of data from two studies of customer satisfaction, in which customer evaluations were measured at several points in time. The customer base (business vs. residential customers), the passivity of the service (local telephone vs. cellular), the service aspect (repair service vs. overall service) and interviewing interval (annual vs. biannual) were different in the studies.

These data afford an analysis of the cognitive mechanism by which service evaluations are created and updated. The language for this analysis comes, variously, from the disciplines of cognitive psychology, marketing science and survey research. From these two datasets, we have found that:

! Reference points, with which current evaluations are contrasted, seem to exist and are heavily influenced by initial or Aprehistoric@ quality evaluations.

! Initial reference points, formed before the survey period, can vary greatly among a customer base and can survive through at least 18 months of service.

! There is no evidence of any quality A inertia@ based on a reference point, which requires an unusual quality change to overcome.

! Current period attributes play a far greater role in service evaluation than any reference point.

It follows that the general level of service evaluation is partly influenced by prior reference values, but this does not preclude incremental rating improvements based on current service

attributes. In other words, aggregate survey results should not breed the fatalism characterized by the belief that opinions are preordained; current (and future) ratings follow service delivery to a very large extent. Instead of their traditional reliance on independent cross-sectional surveys, these results show that the wise company will develop satisfaction measurement programs to follow changes in time series ratings to assess their performance. It is improvement upon goodness, rather than goodness itself, that is most important in quality management.

REFERENCES

- Boulding, W., A. Kalra, R. Staelin and V. Zeithaml (1993), "A Dynamic Process Model of Service Quality: From Expectation to Behavioral Intentions," *Journal of Marketing Research*, XXX (February 1993), 7-27.
- Davis, J. H. (1984), "Order in the Courtroom," in *Perspectives in Psychology and Law*, D. J. Miller, D. G. Blackman and A. J. Chapman (eds.), New York: Wiley.
- Drew, James H. (1990), "Current Findings Regarding BCOS Nonresponse and Stratification," GTE Laboratories Technical Memo.
- Drew, James H. and Ruth N. Bolton (1996), The Dynamics of Customer Opinion, working paper.
- Helson, H. (1964), *Adaptation Level Theory: An Experimental and Systematic Approach to Behavior*, New York: Harper and Row.
- Hogarth, R. and H. Einhorn (1992), "Order Effects in Belief Updating: The Belief-Adjustment Model," *Cognitive Psychology*, 24, 1-55.
- Kahneman, D. and A. Tversky (1973), "On the Psychology of Prediction," *Psychological Review*, 3, 430-454.
- Kalton, G. and C. Citro (1993), "Panel Surveys: Adding the Fourth Dimension," *Survey Methodology*, 19, 2, 205-215.
- LaBarbera, P. A. and D. Mazursky (1983), "A Longitudinal Assessment of Consumer Satisfaction/Dissatisfaction: The Dynamic Aspect of the Cognitive Process," *Journal of Marketing Research*, XX (November 1983), 393-404.
- Nisbett, R. and L. Ross (1980), *Human Inference: Strategies and Shortcomings of Human Judgment*, Englewood Cliffs, NJ: Prentice-Hall.
- Oliver, Richard L. (1980), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, XVII (November), 460-9.
- O'Neill, R. (1971) "Function Minimization Using a Simplex Procedure, Algorithm AS 47," *Applied Statistics*, 338.
- Schul, Y. and M. Schiff (1993), "Measuring Satisfaction with Organizations," *Public Opinion Quarterly*, 57, 4, 536-551.
- Simmons, C., B. Bickart, and J. Lynch, Jr. (1993). "Capturing and Creating Public Opinion in Survey Research," *Journal of Consumer Research*, 20, 2, 316-329.
- Waterton, J. and D. Lievesley (1989), "Evidence of Conditioning Effects in the British Social Attitudes Panel," in *Panel Surveys*, eds. D. Kasprzyk, G. Duncan, G. Kalton and M. Singh, New York: John Wiley and Sons.

Send correspondence regarding this article to:

James H. Drew
GTE Laboratories Inc.
40 Sylvan Road
Waltham, MA 02254 USA

SERVICE QUALITY MEASURES: A TEST OF CONVERGENT VALIDITY AND TRAIT-METHOD EFFECTS

James M. Tolliver, The University of New Brunswick
Robert W. Armstrong, Murdoch University
Daniel F. Coleman, The University of New Brunswick

ABSTRACT

This study reviews and tests seven operationalizations of service quality to ascertain convergent validity. It compares these with earlier parallel operationalizations of job satisfaction, in order to assess the impact of the trait-method units common to both research domains. Results show that the service quality construct has limited convergent validity and further the construct does not explain a major portion of global service quality variance. Comparison of the service quality construct to the job satisfaction construct supplies some evidence of trait-method bias.

INTRODUCTION

"Convergent validity" requires that we have several different measures of the same construct (Campbell & Fiske, 1959). If a particular construct can only be measured by the use of a specific technique or with a single instrument, it raises the possibility that ensuing research results are traceable to the way the construct was measured, rather than to the construct itself. One example of such "method bound" research is the early work of Herzberg and his colleagues (e.g. Herzberg, Mausner, & Snyderman, 1959). This research led to a series of plausible but erroneous conclusions concerning the nature of job satisfaction and motivation. Before the results of this research programme were traced to the methodology employed the results were embraced by practitioners and utilized as the basis of perhaps thousands of company training programmes--programmes that were of doubtful value given the findings of subsequent research.

The requirement that a construct have convergent validity is consequently of interest in any field where researchers report the formulation of specific constructs or practitioners apply such constructs as the basis of management techniques. The area of service quality is no exception. If "service quality" can only be measured in one way or if various measures of service quality do not

converge it raises the possibility that they are "method bound", misleading researchers and practitioners alike. In short, the construct and practice founded on it will be suspect, and subject to the same fate as Herzberg's early work. Fortunately, since 1988, when Parasuraman, Zeithaml and Berry published SERVQUAL, an early instrument designed to measure service quality, the service quality construct has been redefined and operationalized in a number of different ways. Consequently it is now possible to test the convergent validity of several "service quality" measures.

Campbell and Fisk (1959) also note that a construct is a "trait-method unit". Trait-method units are comprised of the stimuli which represent the construct, such as the questions asked on a questionnaire, and the methods employed to analyse the results (e.g. the mathematical methods used to determine a score on the questionnaire). It is also possible the trait-method unit employed to "tap" a particular construct determines the research results. In other words it is possible that any relationship discovered between a construct, such as service quality, and any other construct is due to an interaction between the questions asked and the way in which the answers to the questions are measured, scored, or analyzed. It is therefore useful to examine findings from different, but logically unrelated, research "domains" which employ the same trait-method units to determine if trait-method similarity has an effect on the reported results.

Simply put, any research using a construct such as service quality is suspect if 1) the construct can be measured in only one way (it lacks convergent validity) or 2) if the application of similar questions and similar mathematical techniques in a completely unrelated field gives **exactly** the same pattern of statistical results (indicating a possible trait-method artifact).

This paper consequently takes two approaches to the service quality question. We first evaluate the convergent validity of several measures that purport to measure service quality. This is done by

a traditional correlation analysis and a more appropriate hierarchical regression analysis conducted *within* the service quality domain. Second, we test for a trait-method effect.

The possibility of a trait-method effect is evaluated by comparing the pattern of results from service quality research to the pattern of results from job satisfaction research. Job satisfaction is an unrelated construct, but one which uses the same trait-method units as service quality research. This is a cross domain comparison.

The Trait-Method Units

Service quality and most job satisfaction instruments are comprised of the same trait-method units. In both traditions subjects often indicate, on a series of Likert scales, summary judgments of what **SHOULD BE**, what **IS NOW**, what they **WOULD LIKE** and the **IMPORTANCE** of a variety of items or "facets" concerning a job or a service (For clarity **IS NOW** refers to a job satisfaction construct; the same concept in the service quality domain is indicated by *Is Now*). Such responses are then combine into several dimensions or a general "index". This index is seen as indicative of general job satisfaction or service quality and is sometimes compared to a single global job satisfaction measure or global service quality measure (e.g. Cronin & Taylor, 1992). In some cases difference scores, (or "gap measures") are calculated by subtracting some perceived facet level from some desired level of that facet. Facet *Importance* is also often used as a weight, either for dimensions (e.g. Parasuraman, Zeithaml, & Berry, 1994), or for individual facets when summed as an index (Cronin & Taylor, 1994; Wanous & Lawler, 1972).

In sum, service quality measures and most job satisfaction measures use trait-method units consisting of the same question stems and the same mathematical formulations, often weighted difference scores. It is possible that using **SHOULD BE**, **IS NOW**, **WOULD LIKE** and **IMPORTANCE** stems and then using techniques such as "gap measures" to combine the responses to such stems, or by weighting such responses by **IMPORTANCE**, results in a specific *pattern* of research findings. This pattern might carry across

research areas or "domains".

An examination of the impact of these specific trait-method units across domains has yet to be done, but it is certain that many of the trait-method units described above are problematic. Edwards (1991) identified a number of significant measurement and analytical problems that are traceable to these trait-method units in a body of literature he calls the "Person-Job", or "P-J Fit" literature. He concludes that the data collection and analytical methods currently used in the Person-Job literature, including job satisfaction, could well be misleading (Edwards, 1991). The literature Edwards specifically criticizes views work outcomes (such as job satisfaction) as a match between "desires" (such as needs, values, or preferences) and "supplies" (such as job attributes or workload). This is, in effect, the same analytical form ("gap" analysis) that is found in a lot of service quality research. Edwards (1991), however, confined his review to the P-J literature, including job satisfaction, leaving open the possibility that similar problems *across* domains--specifically in service quality--could be due to the use of such trait-method units.

The Job Satisfaction Construct. In 1972 Wanous and Lawler identified nine trait-method units that are prevalent in the job satisfaction literature. These trait-method units, presented in Table 1, still clearly define conceptualization and measurement in the job satisfaction field.

While the formulae presented in Table 1 are rather direct it is useful to briefly provide definitions of the terms used: **MEAN OF FACET SATISFACTION** is an average satisfaction score calculated by directly asking each subject "how satisfied are you with x?" (where x is a specific item or facet such as pay). **IS NOW** asks how much of a facet, such as pay, is present. It is a measure of goal attainment or need fulfilment (Wanous & Lawler, 1972). **IMPORTANCE** is a rating of a facet's importance. **SHOULD BE** is a "fairness" construct; specifically, it is the subjects' judgement of "how much" of a facet, such as pay, is "fair" or equitable given the input required of the job. Finally, **WOULD LIKE** is a measure of the desired or "ideal" levels of an outcome.

IS NOW, **IMPORTANCE**, **SHOULD BE**, and **WOULD LIKE** are usually rated on a series of

Table 1
Operational Definitions of Overall Satisfaction
Correlations with Single Item Measuring Overall Satisfaction^a

Equation Numbers	Absolute Value Correlation (r)	Measure ^b
1	.61	Is Now
2	.60	Mean of Facet Satisfaction
3	.54	Would Like - Is Now
4	.50	Importance x Facet Satisfaction
5	.48	Importance x Is Now
6	.45	Importance x (Would Like - Is Now)
7	.44	Importance - Is Now
8	.39	Importance x (Should Be - Is Now)
9	.24	Should Be - Is Now

^a From Wanous and Lawler (1972), Table 1, page 98. Our equation numbers do *not* parallel Wanous and Lawlers' since we have chosen to present their findings as a partial replication of this original table.

^b This column contains a description of the measures in words. The mathematical representation of each measure can be found in Table 2.

seven point Likert scales that range from "a lot" to "a little" or, in the case of IMPORTANCE, from "very important" to "not very important". Subjects respond to these "stems" for specific facets such as pay, promotional opportunities, and relations with co-workers. Typically a job satisfaction questionnaire might ask a respondent how much pay there IS NOW or how much pay there SHOULD BE and also ask the subject to rate the IMPORTANCE of pay. While different job satisfaction questionnaires use different stems and different mathematical combinations of them, these different combinations can be classified as indicating overall satisfaction (equations 2 and 4, Table 1), indicating fulfilment (1 and 5, Table 1), indicating perceived equity (8 and 9, Table 1), and as desired or value models (3 and 6, Table 1) (Equation 7 is seen as "theoretically meaningless") (Wanous & Lawler, 1972; Evans, 1969).

Wanous and Lawler reasoned that if such a

variety of different trait-method units were tapping a common construct, and the influence of this common construct was greater than the unique variance tapped by the measures, the responses to these different trait-method units should be highly related. In other words, if a general "job satisfaction" factor was being tapped by these different measures they should be highly correlated. They tested the convergent validity of these nine trait-method units by using a twenty-three facet instrument that focused on such items as pay and supervision. The nine trait-method units represented by the equations in Table 1 failed the test of convergent validity. IS NOW (measure 1, Table 1) correlated most highly with a global satisfaction measure while the equity measure (measure 9, Table 1) achieved the lowest correlation. IMPORTANCE was not useful as a facet weight but was related to the impact a particular facet had on global satisfaction (Wanous

& Lawler, 1972). In other words these trait-method units had little in common and were not similarly related to a global measure of satisfaction. This raised the possibility that incompatible research findings were due to the use of measures that claimed to measure the same construct but in fact did not. And while not stated directly, a failure to find convergent validity raised the possibility that the job satisfaction construct could not be measured adequately at all.

By implication the same conclusions might be applied to service quality research since, as we show below, current service quality measures use exactly the same trait-method units which lacked convergent validity in the job satisfaction literature.

The Service Quality Construct. SERVQUAL (Parasuraman, Zeithaml, & Berry, 1988), is a twenty-two facet instrument that taps "service quality". The original SERVQUAL asked subjects to indicate what service SHOULD BE offered by a firm (denoted as "*Expectations*" (E)) and the *Perception* (P) that a target firm had the described facet. Service quality [Q], a gap measure, is defined as $Q = P - E$. The results are summed across facets to represent a total measure of service quality.

Subsequently Carman (1990) concluded that facet *Importance* (IMPORTANCE) may be more relevant than *Expectations* (SHOULD BE). Babakus and Boller (1992) questioned the usefulness of gap analysis, concluded that the SERVQUAL may be uni-dimensional, and tested *Importance* as a facet weight. They discovered *Importance* added little to the analysis.

By using a SHOULD BE stem and the perception (IS NOW) stem, Parasuraman, Zeithaml, and Berry, (1988), used a trait-method unit directly comparable to the "Equity" measure of job satisfaction (SHOULD BE - IS NOW, measure 9, Table 1) while Babakus and Boller, (1992) are, in effect, using those trait-method units denoted as 1, 5, 8, and 9 in Table 1. Babakus and Boller's findings (1992) parallel an earlier discovery in the job satisfaction literature that using IMPORTANCE as a weight is an ineffective technique (e.g. Blood, 1971; Evans, 1969). Similar to Wanous and Lawler's finding that IS NOW explained the most variance in a

global measure of job satisfaction, SERVPERF (an IS NOW measure) explained more variation in a global measure of service quality than the SERVQUAL. Others (Teas, 1993) raise the issue of "Ideal Points" (a WOULD LIKE trait-method unit), and also question the dimensionality of the original and revised SERVQUAL models (Teas, 1994).

From this admittedly brief review it is clear that the trait-method units used in service quality research and job satisfaction research (as well as some other areas of the Person-Job fit literature) are the same. The extent of this similarity can be seen from Table 2 which pairs service quality trait-method units with comparable job satisfaction trait-method units.

The fact that the trait-method units used in service quality research are the same as those used in job satisfaction research presents an opportunity to evaluate the impact of the various trait-method units used. While convergent validity can be determined by comparing the results from using different service quality trait-methods--a within domain analysis--a cross-domain analysis can be conducted by comparing the pattern of results achieved using various service quality trait-method units to the results achieved with the same trait-method units used in job satisfaction. Consequently in the analysis below we have included gap measures which are still much in vogue (Oliver, 1997) and also included a measure of *Importance*. While we are aware of the current debate (e.g. Oliver, 1997) concerning the concepts of service quality and service satisfaction, we have included those trait-method units which various authors (e.g. Parasuraman, Zeithaml & Berry, 1988) have claimed to be service quality measures. This permits us to directly compare findings across domains and to ascertain if the limited impact of some variables, such as *Importance*, is due to an inappropriate analysis. Given differences in the domains, the samples, the cultures, the facets and the time lines (the data describe below were collected more than 20 years after the job satisfaction data used for comparison) it is likely that a similar result pattern is due to method variance traceable to the trait-method units employed rather than to similarities in the samples used. Thus, this paper addresses two research questions:

Table 2
Wanous and Lawler Job Satisfaction Operationalizations
Compared with Possible Operationalizations of Service Quality

NUMBERS FROM TABLE 1	WANOUS/LAWLER DESCRIPTION	EQUATION	SERVICE QUAL. EQUIVALENT	EQUATION
1	Is Now	$JS = \Sigma (\text{Is Now})$	Perceptions	$Q = \Sigma (P_i)$
3	Would Like - Is Now	$JS = \Sigma (\text{Would Like} - \text{Is Now})$	Ideal - Perceptions	$Q = \Sigma (I_i - P_i)$
5	Importance x Is Now	$JS = \Sigma (\text{Importance} * \text{Is Now})$	Importance x Perceptions	$Q = \Sigma (I_i * P_i)$
6	Importance x (Would like - Is Now)	$JS = \Sigma [\text{Importance} (\text{Would Like} - \text{Is Now})]$	Importance x (Ideal - Perceptions)	$Q = \Sigma I_i (I_i - P_i)$
7	Importance - Is Now	$JS = \Sigma (\text{Importance} - \text{Is Now})$	Importance - Perceptions	$Q = \Sigma (I_i - P_i)$
8	Importance x (Should Be - Is Now)	$JS = \Sigma [\text{Importance} (\text{Should Be} - \text{Is Now})]$	Importance x (Expectations - Perceptions)	$Q = \Sigma I_i (E_i - P_i)$
9	Should Be - Is Now	$JS = \Sigma (\text{Should Be} - \text{Is Now})$	Expectations - Perceptions	$Q = \Sigma (E_i - P_i)$

Research Question 1. Convergent Validity. What is the relationship between the different trait-method units of service quality and how are these operationalizations related to the global measure of service quality? If the measures of service quality have convergent validity we expect the various trait-method units to be highly correlated with each other and to be highly and uniformly correlated to a global measure of service quality.

Research Question 2. Trait-method effects. If the same trait-method units are used in different research domains are the results the same? Specifically, is the **pattern** of correlations between the trait-method units used in service quality research similar to the **pattern** of correlations between the trait-method units used in job satisfaction research? If there is a strong trait-method effect we would expect the pattern of correlations found between the service quality trait-method units to be the same as the pattern of relations between the job satisfaction trait-method units.

THE CURRENT STUDY

Sample

Customers and clients of over twenty different Australian organisations in fifteen different service industries were sampled. A total of 3000 questionnaires were mailed or hand delivered; 1135 subjects (38%) responded comprising a final sample of 18 different organisations across the fifteen different industries. In most cases a questionnaire was given to the nth person sitting in an airline terminal, restaurant, medical centre or bank queue. In other industries, random sampling methods taken from credit card listings were used. MBA students were also sampled via convenience method. In most cases, probability sampling methods were used. In some cases proprietors offered incentives to complete the surveys (e.g., discount coupons, free drinks, etc.). Consequently, the method of data collection was somewhat different for some of our subjects (i.e. mail or a hand delivered questionnaire). A test of the major variables, however, did not find any

difference due to the data collection method used.

The median number of responses within each industry is approximately 100. The respondents' average age is 29.81 years. There are 445 male and 620 female respondents; 70 persons did not indicate gender. The subjects are approximately evenly distributed across industries. The mean level of previous experience with all industries combined is 1.21 years with a minimum of zero (no experience) to a maximum 60 years.

Measurement Instrument

The original 22-item SERVQUAL scale developed to measure service *Perceptions* and *Expectations* was rewritten to include service *Importance* and *Ideal* service stems. The negatively worded questions (a problem in the SERVQUAL; Babakus and Boller, 1992) were rewritten in the positive. *Importance* was measured on each of the 22 Parasuraman, Zeithaml, and Berry, (1988) facets using a 7 point-scale (1 "Not Important"; 7 "Very Important"). *Ideal* service was measured for each facet by asking respondents to indicate "How much of an attribute should be present in an *Ideal* organization" (1 "None"; 7 "A Great Deal"). The *Ideal* scale differs from the *Expectations* scale in that the *Expectations* scale is based on the question, "Please show the extent to which you think XYZ should possess the feature described in each statement." In short, the *Ideal* scale asks respondents to imagine what they WOULD LIKE while the *Expectation* scale asks respondents to indicate what the service SHOULD BE for a particular firm. This wording attempts to directly tap the "ideal" point discussed in the literature (Parasuraman, Zeithaml, & Berry, 1994; Teas, 1993; Teas, 1994).

In addition to the 22 item SERVQUAL scale, a seven point (1 Low; 7 High) global measure of service quality, "Overall how would you rate the service quality of XYZ?" was included.

Analyses

Research Question 1. Convergent Validity. The relationship between the seven trait-method units used in service quality research may be analysed in two ways. The first method is to perform the calculations required by the various

formulations contained in Table 2 then a) correlate these results with the global measure of service quality and b) determine the inter-correlations between the seven service quality operationalizations. This is the "analysis by correlation" method originally used by Wanous and Lawler (1972), recommended by Campbell and Fiske (1959), and used in a number of service quality studies. While it is a weaker analytic approach it is included for the purposes of comparison to earlier work.

The second, stronger, method uses hierarchical regression to enter the component parts of those trait-method units that employ difference scores separately -- a procedure which is more acceptable than using the correlation procedure. The reasoning for this "composite procedure" can be found in Schmidt and Wilson (1975), in Edwards and Cooper (1990) and in Evans (1991) and is explained more fully below.

RESULTS: QUESTION 1

Correlational Analyses

Table 3A contains the correlations of each service quality trait-method unit with the other service quality trait-method units. The correlation of the various service quality trait-method units with the global measure of service quality are compared in the last three columns of Table 3A. Table 3A requires that the sometimes "weighted" difference scores for each subject be summed across all 22 SERVQUAL items before the correlations are computed. While the content, number and stability of the dimensions which comprise the SERVQUAL has been the subject of debate (e.g. Carman, 1990), Cronin and Taylor (1994) contend that SERVQUAL is best represented by a single sum across all items.

The expectation of convergent validity is that measures be highly related to each other and to the global measure of service quality. While some trait-method units are related arguing for convergent validity, Table 3A shows that significant differences do exist. In particular, the relationships of the *Ideal* or *Expectations* trait-method units with other trait-method units are different. For example, the correlations between the *Ideal* (Id) trait-method units and *Perceptions*

Table 3A
Service Quality Equations Correlation Matrix^a

Equation Number From Table 1		Service Quality Description	$\Sigma (P_i)$	$\Sigma (Id_i - P_i)$	$\Sigma (I_i * P_i)$	$\Sigma I_i (Id_i - P_i)$	$\Sigma (I_i - P_i)$	$\Sigma I_i (E_i - P_i)$	$\Sigma (E_i - P_i)$	Correlations with Global Measure ^b		
										This Study	C & T ^c	B & B ^d
1	$\Sigma (P_i)$	Perceptions	1	0.60	.90 (.91)	0.59	0.77	.82 (.80)	.82 (.81)	0.77	0.60	0.66
3	$\Sigma (Id_i - P_i)$	Ideal-Perceptions		1	0.42	0.99	0.66	0.64	0.64	0.48	--	--
5	$\Sigma (I_i * P_i)$	Importance x Perceptions			1	0.41	0.44	.57 (.63)	.55 (.66)	0.67	0.56	--
6	$\Sigma I_i (Id_i - P_i)$	Importance x (Ideal - Perceptions)				1	0.67	0.65	0.64	0.47	--	--
7	$\Sigma (I_i - P_i)$	Importance - Perceptions					1	0.91	0.92	0.64	--	--
8	$\Sigma I_i (E_i - P_i)$	Importance x (Expectations - Perceptions)						1	.99 (.98)	0.69	0.54	--
9	$\Sigma (E_i - P_i)$	Expectations - Perceptions							1	0.70	0.54	0.59

^a All correlation coefficients were significant at the 0.0001 level.

^b All correlations were significant at the 0.0001 level.

^c Cronin and Taylor (1992).

^d Babakus and Boller (1992).

(IS NOW) trait-method units is significantly smaller than the relations between the *Expectations* trait-method units and *Perceptions*. All such comparisons are statistically significant ($p = .001$, or beyond, Fisher r to Z transformation). Another trend is that the *Ideal* trait-method units are closely related to each other, but are not as strongly related to the *Expectation* units. Finally, replicating previous research, all correlations remain relatively unchanged if *Importance* (I_i) is present or absent.

The correlations reported by Cronin and Taylor (1992)--who use this method--are reported in parentheses for comparison purposes. The relative magnitude of the correlations discovered in the Australian sample replicate the relative magnitude of those correlations reported by Cronin and Taylor (1992) with a United States sample

(Kendall's Tau, $\lambda = .83$) indicating apparent cross-cultural stability.

Table 3A also reports correlations of the various measures with a single global measure of service quality. The correlations from Babakus and Boller (1992) and, again, Cronin and Taylor (1992) are included as comparisons. In Table 3A, the summation of the *Perceptions* ($Q = \Sigma(P_i)$) is most strongly related to a global measure of service quality (.77, $p < .001$). The next highest correlation (.70) is the trait-method unit recommended by Parasuraman, Zeithaml, and Berry (1988) [$Q = \Sigma(E_i - P_i)$]. *Importance* (I_i) does not improve the correlation with the global measure. The *Ideal* (Id_i) trait-method unit does not improve the explained variance of the global measure (correlations .48 and .47). This pattern extends the findings of Cronin and Taylor (1992)

and Babakus and Boller (1992) concerning *Perceptions* ($Q = \Sigma(P_i)$).

Regression Results

While the "analysis by correlation" method has been used extensively, and was conducted in this paper for the purposes of comparison with earlier research, a more appropriate test of convergent validity is accomplished via hierarchical regression. As Teas (1994) notes the calculation of differences may result in sums that are at or near zero. This was the case for the *Should Be-Is Now*, *Would Like-Is Now* and *Importance-Is Now* data in this study and, as noted by one reviewer, further limits the usefulness of the "analysis by correlation" method. But the correlational approach is also questionable on theoretical rather than empirical grounds.

Schmidt and Wilson (1975) recommend hierarchical regression because the multiplication of two Likert scales can result in multiplying terms that contain deviations from "true" zero. For example, weighting *Perceptions* (P_i), by *Importance* (I_i) (to construct a trait-method unit such as equation 5 in Table 3A) is a multiplication of a score and its deviation from true zero (if any) with another score and its deviation from true zero. In Schmidt and Wilson's (1975) terms, the measure of *Perceptions* has a "true" component (P_i) and some unknown error component (e_p). The *Perceptions* measure is then actually $P_i + e_p$, and the *Importance* measure is similarly $I_i + e_i$. The multiplicative term of $I_i * P_i$ is actually $I_i P_i + I_i e_p + P_i e_i + e_i e_p$. Consequently, if a transformation is applied to I_i , to P_i or both, say by rescaling the data, this legitimate transformation could result in a substantial change in the correlation of $I_i * P_i$ with the global measure of service quality (or any other measure, even other composite operationalizations) (see Schmidt & Wilson, 1975). Since convergent validity is dependant on trait-method correlations, rescaling effects the conclusion of convergent validity or the lack of it.

Evans' (1991) suggestion is to consider any multiplicative form (such as many of the trait-method units in Table 1) as an interactive term in a linear regression, and to analyse any relationship involving such a term by first taking into account its component parts as main effects. This requires

that those terms used for multiplication be first entered in a regression as "main effects" and that the impact of the multiplicative term subsequently be assessed. For example, in evaluating the relationship of $I_i * P_i$ to another variable, the variables that comprise $I_i * P_i$ (I_i , and P_i respectively) must be entered into a regression **first**. The test of any multiplicative trait-method unit's relationship with any other trait-method unit is then the ability of the interaction ($I_i * P_i$) to add significantly to the R^2 after the component parts, or "main effects" that comprise the interaction have been entered. This technique is preferred to the correlation approach although it "...relies on the assumption of linearity between the underlying psychological variables and the measures used." (Evans, 1991: 7).

Edwards and Cooper (1990) argue that terms which comprise difference (i.e. gap) measures, such as many of the trait method units in Table 1, must also be entered separately into the equation. This is so because **any** operationalization which uses a gap measure, even an absolute value, in the form of $A_i = a_0 + b_1 (E_i - P_i) + e$ is a more restrictive, misleading, form of the equation $A_i = a_0 + b_1 (E_i) - b_2 (P_i) + e$ (For example, in the case of the SERVQUAL: A_i could be global satisfaction; a_0 the intercept, b 's are conventional beta weights, P_i is the *Perceptions* measure and E_i is the *Expectations* measure. The error term is e .)

The restrictive model effectively **requires** the beta weights for both P_i and E_i to be the same and also **requires** that P_i be subtracted from E_i . When using the second model, however, these constraints **will** be met only if a "gap" measure is truly a good predictor of the dependant variable. The resulting equation **will** have beta weights approximately the same and the sign of the second beta weight **will** be negative due to the characteristics of the data. In the case of the more appropriate second model, however, the equation will arise from the data and not have been forced by the trait-method unit used. Consequently an equation such as 9 (Table 2) would be $Q = a_0 + b_1(E_i) - b_2(P_i) + e$.

To test the effect of these restrictions and to meet the objections outlined above we conducted two regressions using the trait-method units of service quality as independent variables and the global service quality measure as the dependant

Table 3B
Hierarchical Regression Results

Equation with Global Service Quality as dependent variable	R ²	Degrees of Freedom	F-Score
$\Sigma(P_i)$	0.60	919	1384.24
$\Sigma(I_i)$	0.60	919	521.45
$\Sigma(P_i) + \Sigma(Id_i - P_i)$	0.60	897	672.16
$\Sigma(P_i) + \Sigma(Id_i - P_i) + \Sigma(I_i * P_i)$	0.60	865	434.56
$\Sigma(P_i) + \Sigma(Id_i - P_i) + \Sigma(I_i * P_i) + \Sigma I_i(Id_i - P_i)$	0.60	865	326.01
$\Sigma(P_i) + \Sigma(Id_i - P_i) + \Sigma(I_i * P_i) + \Sigma I_i(Id_i - P_i) + \Sigma(I_i - P_i)$	0.60	865	266.09
$\Sigma(P_i) + \Sigma(Id_i - P_i) + \Sigma(I_i * P_i) + \Sigma I_i(Id_i - P_i) + \Sigma(I_i - P_i) + \Sigma I_i(E_i - P_i)$	0.61	850	221.64
$\Sigma(P_i) + \Sigma(Id_i - P_i) + \Sigma(I_i * P_i) + \Sigma I_i(Id_i - P_i) + \Sigma(I_i - P_i) + \Sigma I_i(E_i - P_i) + \Sigma(E_i - P_i)$	0.61	850	190.67

Table 3C
Hierarchical Regression: Edwards and Cooper Method

Equation with Global Service Quality as dependent variable	R ²	Degrees of Freedom	F-Score
$\Sigma(P_i)$	0.60	919	1384.2
$\Sigma(I_i)$	0.60	919	521.45
$\Sigma(P_i) + \Sigma(E_i)$	0.60	883	687.33
$\Sigma(P_i) + \Sigma(E_i) + \Sigma(I_i)$	0.60	870	451.97
$\Sigma(P_i) + \Sigma(E_i) + \Sigma(I_i) + \Sigma(Id_i)$	0.60	865	437.56
$\Sigma(P_i) + \Sigma(E_i) + \Sigma(I_i) + \Sigma(Id_i) + \Sigma(I_i * P_i)$	0.60	865	338.65
$\Sigma(P_i) + \Sigma(E_i) + \Sigma(I_i) + \Sigma(Id_i) + \Sigma(I_i * P_i) + \Sigma(I_i * E_i)$	0.60	865	270.61
$\Sigma(P_i) + \Sigma(E_i) + \Sigma(I_i) + \Sigma(Id_i) + \Sigma(I_i * P_i) + \Sigma(I_i * E_i) + \Sigma(I_i * Id_i)$	0.60	850	220.02

Table 4
A Comparison of the Correlations from the Wanous and Lawler
Equations and Service Quality Equations

Equation Number From Table I	Wanous and Lawler Equation	Service Quality Equivalent	Wanous & Lawler Correlation	Service Quality Correlation
1	$JS = \Sigma (Is\ Now)$	$Q = \Sigma (P_i)$	0.61	0.77
3	$JS = \Sigma (\text{Would Like} - Is\ Now)$	$Q = \Sigma (Id_i - P_i)$	0.54	0.48
5	$JS = \Sigma (\text{Importance} * Is\ Now)$	$Q = \Sigma (I_i * P_i)$	0.48	0.67
6	$JS = \Sigma [\text{Importance} (\text{Would Like} - Is\ Now)]$	$Q = \Sigma I_i (Id_i - P_i)$	0.45	0.47
7	$JS = \Sigma (\text{Importance} - Is\ Now)$	$Q = \Sigma (I_i - P_i)$	0.44	0.64
8	$JS = \Sigma [\text{Importance} (\text{Should be} - Is\ Now)]$	$Q = \Sigma I_i (E_i - P_i)$	0.39	0.69
9	$JS = \Sigma (\text{Should Be} - Is\ Now)$	$Q = \Sigma (E_i - P_i)$	0.24	0.70

variable. The global service quality measure was selected since service quality is seen to be a global concept. The results of these regressions are reported in Table 3B and Table 3C. Since in our analysis it is the strength of the relationship not its form that is of concern we report the R squares but not the beta weights.

Table 3B reports the results when the trait-method units, including the results of the gap measures, are regressed directly on global service quality. Since the components used to calculate the difference scores are **not** entered separately in Table 3B we can compare the results of the "gap" models as currently conceptualized to the results we obtain when the objections of Evans (1991) and Edwards and Cooper (1990) are met. The second regression utilized the components separately as recommended by Evans and others. Both regressions were conducted in stages for the reasons explained above.

The improvement of R^2 in table 3B shows that the summed *Perceptions* measure ($Q = \Sigma(P_i)$) is the best predictor of global service quality. The other composite operationalizations do not enter the regression. This means that the other trait-method units have nothing to add to explaining global service quality. This finding supports the correlation results (Table 2) and the research of Cronin and Taylor (1992).

It may have been possible that the results in Table 3B are due to the restrictions placed on the trait-method units by **first** calculating the difference scores and then utilizing these scores in the regression, this is a likely interpretation given the problem outlined by Teas (1994) and noted earlier. Consequently Table 3C reports the results when the *components* of the composite operationalizations are entered separately (Evans, 1991; Edwards and Cooper, 1990).

In this case the simple *Perceptions* measure is again most closely related to the global measure of service quality.

All analyses--correlations, and both regressions--give similar results. The *Perceptions* measure is more closely related to global perceptions of service quality than any other trait method unit and other formulations add nothing to the analysis.

RESULTS: QUESTION 2

Research Question 2. Trait-method effects. The cross-domain effect is tested by an examination of the **pattern** of correlations between the trait-method units and the global measure of job satisfaction or service quality respectively. If there is a trait-method effect we would expect to see the pattern of correlations found in job satisfaction to be replicated in the service quality data. (The job satisfaction correlations are taken from Table 1 and the service quality correlations from Table 3A).

The correlations in Table 4 were compared using two methods. First a Spearman rank correlation test was run treating each trait-method unit as an "object" and the correlations associated with each trait-method unit as a "pair" of observations. If there was a series of trait-method effects those trait-method units that correlate highly with a global measure of job satisfaction would be those same trait-method units that correlate highly with the global measure of service quality. The effect within each trait-method unit was examined by transforming the correlations to Z scores and comparing them. These tests reveal that service quality operationalizations yield different coefficients in terms of both pattern and magnitude than the job satisfaction trait-method units used as comparisons. The rank correlation test was not significant, and, in 5 of the 7 cases, the sizes of the correlations were significantly different across research domains ($p = .001$).

The service quality *Expectations* (SHOULD BE) is a much better predictor of global service quality than SHOULD BE was at predicting global job satisfaction. In both equation sets, *Importance* (IMPORTANCE) did not improve the explained variance. *Perception* (IS NOW) was the best predictor of global service quality and global job

satisfaction but the correlations are of different magnitude ($p = .001$).

DISCUSSION

This paper uses proposals by Edwards (1991), Evans (1991), and others to test the convergent validity of different trait-method units of service quality. It also compares the **pattern** of correlations between different service quality trait-method units and a global measure of service quality to the **pattern** of correlations between the same trait-method units used in job satisfaction research and a global job satisfaction measure (Wanous & Lawler, 1972). The purpose of this comparison is to determine if the same trait-method units result in a common pattern of results across two different domains--a possibility given the findings of Edwards (1991) and the objections of Edwards and Cooper (1990).

Findings show that some service quality trait-method units are weakly correlated to others. Further, some composite operationalizations are differentially correlated to global service quality with *Perceptions* (an IS NOW measure) most strongly related to the global measure of service quality. This is supported using hierarchical regression as well as the more traditional, but flawed, "analysis by correlation" method.

While some service quality trait-method units are related--making a reasonable argument for convergent validity --this result must be approached with caution. From the data it seems that the results are due to the commonality of a general perceptual "set" tapped by *Perceptions* (IS NOW). This might be encouraging for service quality researchers, yet the use of *Perceptions* as the trait-method unit of choice is a problem since much of the variance in the global measure of service quality still remains unexplained. This is significant since service quality is a summary judgement about overall excellence or superiority. It is also problematic because "Validity is represented in the agreement between two attempts to measure the same trait through maximally different methods" (Campbell & Fiske, 1959: 83). In short validation typically requires confirmation by measurement procedures that are independent and maximally different but related. If this can not be shown, the construct under study is

questionable on the grounds that it can, essentially, be measured in "only one way"--or, in the present study, effectively taped only by an IS NOW approach.

Finally, as others have noted, the concept of a number of trait-method units (e.g. "fit" or gap measures; Edwards, 1991) are ubiquitous in psychology and related disciplines. In fact a large number of approaches in different research domains rest on a surprisingly small number of trait-method units such as those found in both the job satisfaction and service quality literatures just reviewed. In such instances it is useful to determine if the "findings" seen as unique in one domain are due to the phenomenon under study or due to the method of data collection and analysis, that is the trait-method unit, used. This paper attempts to provide such an answer in the field of service quality. This interpretation -- that patterns of correlations are being repeated across domains due to the trait-method unit used -- receives marginal support at best. The rank-order of correlations found by Wanous and Lawler (1972) between the trait-method units used and a global measure of job satisfaction is not repeated when the same trait-method units are used in service quality research. An important exception to the dissimilar results, however, is the fact that the *Perceptions* (IS NOW) facets, at the core of the responses as noted above, were more strongly correlated with the global measures of both service quality and satisfaction than any other single measure or composite operationalization of either concept. It could therefore be the case that there is a method bias in both areas--that the IS NOW perceptual set dominates perceptions both in the service quality and job satisfaction domains. This might further be the case if similar facets (or "roots") were examined. In this study the service quality facets used were, with the exception of rewording of negative items, taken from the SERVQUAL and were chosen to make this study comparable to other service quality work. The facets in the Wanous and Lawler (1972) work dealt not with quality issues but with items such as pay and promotion opportunities. To the degree the facets or "roots" are considered as part of the trait-method unit (e.g. Oliver, 1997) this approach reduces the possibility of significant findings since the facets used in the two domains are not

common. Further the Wanous and Lawler (1972) work was conducted over 20 years earlier and in the United States. The current data is therefore separated from the Wanous and Lawler comparison data not only by culture (an Australian sample) but by a significant amount of time and, perhaps, social change. Consequently, finding any pattern similarities becomes more difficult and the discovery of a significant "IS NOW" effect raises the possibility that the trait-method problem is more severe than the current data would indicate.

Considered together these findings lead to the conclusion that there is some marginal evidence for a trait-method effect, and that convergent validity may be due to the presence of a single common "set" taped by primarily by an *Perceptions* (IS NOW) trait-method unit. This paper, however, does not graph the form of the "raw" service quality data generated against forms generated by different "fit" indices such as the IS NOW model. When Edwards (1991) conducted similar tests he concluded that the data generated could "...take on a variety of forms, ranging from a simple sloped plane to a complex curvilinear surface." (Edwards, 1991: 346). He further discovered that while different fit models might produce statistically significant results they could misrepresent the raw data form. A similar objection, but on very different grounds, is raised by Oliver (1997) and others with the observation that expectancy disconfirmation might be more adequately represented by an "s" curve. Our paper asked if misrepresentations across research domains are due to the trait-method unit used (of which linear analytic tradition is a part) but did not investigate the suitability of the *Perceptions form* within the Service Quality research domain. This we leave for another paper.

REFERENCES

- Babakus, E. and G. W. Boller (1992), "An Empirical Assessment of the SERVQUAL Scale," *Journal of Business Research*, 24, 253-268.
- Blood, M. R. (1971), "The Validity of Importance," *Journal of Applied Psychology*, 55, 487-488.
- Campbell, D. T. and D. W. Fiske (1959), "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," *Psychological Bulletin*, 56, (2), 81-105.

- Carman, J. M. (1990), "Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions," *Journal of Retailing*, 66, (1), 33-55.
- Cronin, J. J. Jr. and S. A. Taylor (1992), "Measuring Service Quality: A Reexamination and Extension," *Journal of Marketing*, 56, 55-68.
- Cronin, J. J. and S. A. Taylor (1994), "SERVPERF Versus SERVQUAL: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality," *Journal of Marketing*, 5, 125-131.
- Edwards, J. R. (1991), "Person-Job Fit: A Conceptual Integration, Literature Review, and Methodological Critique," *International Review of Industrial and Organizational Psychology*, C. L. Cooper and I. T. Robertson, (eds.) New York: Wiley, 6, 283-357.
- Edwards, J. R. and Cooper, C. L. (1990), "The Person-Environment Fit Approach to Stress: Recurring Problems and Some Suggested Solutions," *Journal of Organizational Behaviour*, 11, 293-307.
- Evans, M. G. (1969), "Conceptual and Operational Problems in the Measurement of Various Aspects of Job Satisfaction," *Journal of Applied Psychology*, 53, 93-101.
- Evans, M. G. (1991), "The Problem of Analysing Multiplicative Composites: Interactions Revisited," *American Psychologist*, 46, (1), 6-15.
- Herzberg, F., B. Mausner and Snyderman (1959), *The Motivation to Work*, New York: Wiley.
- Parasuraman, A., V. A. Zeithaml L. L. and Berry (1988), "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perception of Service Quality," *Journal of Retailing*, 64, (1), 12-40.
- Parasuraman, A., V. A. Zeithaml and L. L. Berry (1994), "Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research," *Journal of Marketing*, 58, 111-124.
- Schmidt, F. L. and T. C. Wilson (1975), "Expectancy Value Models of Attitude Measurement: A Measurement Problem," *Journal of Marketing Research*, 12, 366-368.
- Teas, R. K. (1994), "Expectations as a Comparison Standard in Measuring Service Quality: An Assessment of a Reassessment," *Journal of Marketing* 58, 132-139.
- Teas, R. K. (1993), "Expectations, Performance Evaluation, and Consumers' Perceptions of Quality," *Journal of Marketing*, 57, 18-34.
- Wanous, J. P. and E. E. Lawler, III (1972), "Measurement and Meaning of Job Satisfaction," *Journal of Applied Psychology*, 56, (2), 95-105.

Correspondence regarding this article should be sent to:

James M. Tolliver
The University of New Brunswick
Post Office Box 4400
Fredericton New Brunswick
CANADA E3B 5A3

COMPARISON STANDARDS IN CONSUMER SATISFACTION FORMATION: INVOLVEMENT AND PRODUCT EXPERIENCE AS POTENTIAL MODERATORS

Jong-Won Park, Korea University
Jiho Choi, Doctoral Student, Korea University

ABSTRACT

The extant literature suggests that consumers may use one or some of different comparison standards to evaluate actual product performance during consumer satisfaction formation. This research intended to identify boundary conditions under which a particular standard operates. Four alternative standards were considered: expectation, norm, ideal, and equity. These were comparatively tested by analyzing the data collected over various consumption situations. Two major findings emerged. First, consumer involvement and product experience appeared to interactively influence the type of comparison standard used for evaluating product performance. Second, consumers did not use multiple standards simultaneously in any situations. Theoretical and managerial implications of these findings are discussed.

BACKGROUND

Consumer satisfaction has long been investigated by marketing and consumer researchers. In particular, a number of empirical studies have been published with regard to cognitive processes mediating consumer satisfaction (see Yi 1990 for a review). In these studies is typically employed a form of "standard-performance-disconfirmation" paradigm: (1) actual product performance is compared to an internal standard during satisfaction formation, and (2) the level of the internal standard, the level of actual product performance, and/or the magnitude of the difference between the two levels are determinants of a consumer satisfaction level. The literature suggests that the type of internal comparison standard is diverse. In some cases, for example, a brand expectation (the level at which the brand performance is expected to be) is such a standard (Oliver 1980; Oliver and Linda 1981; Westbrook and Reilly 1983), and in other cases a product norm (the level at which a consumer feels the product performance should be (Cadotte,

Woodruff, and Jenkins 1987), ideal (the level at which the product performance is personally desired to be (Tse and Wilton 1988), or equity (the product performance level which is considered equitable to the price paid or effort invested (Oliver and DeSarbo 1988; Tse and Wilton 1988). Further, recent studies reported that more than one standard might be used simultaneously (Tse and Wilton 1988; Spreng, MacKenzie, and Olshavsky 1996). Then, it seems reasonable to suppose that the type of internal standards used might depend upon some contextual conditions. Identifying such conditions will increase our understanding of more precise cognitive mechanisms mediating consumer satisfaction and provide more meaningful managerial implications. Unfortunately, no empirical investigation has yet dealt with this important issue. The results reported here provide some insights into this matter.

Types of Comparison Standards

Perhaps, the work by Oliver (1980) is one of the pioneering studies formally looking into cognitive processes underlying consumer satisfaction. In that study, he proposed the "expectation-disconfirmation" paradigm, in which the prepurchase "expectation" of a product served as a comparison standard in the satisfaction formation process. Following this work, a great deal of research effort has been devoted to test and extend it. In the meantime, researchers also began to consider alternative comparison standards such as product norms, ideal, and equity. Following is a brief description of these alternative standards and of related research results. Since a much richer review is available elsewhere (Yi 1990), our description here will be brief.

Expectation. This standard is most frequently employed in satisfaction studies (Bearden and Teel 1983; Churchill and Surprenant 1982; Day 1984; Oliver 1980; Oliver and Linda 1981; Swan and Trawick 1981; Westbrook and Reilly 1983).

Expectation is a prepurchase cognition about how good the product performance **will** be and is typically measured using a bipolar semantic differential scale such as "how good or bad did you expect the product performance would be."

Product Norm. Consumers may evaluate actual product performance against some sort of normative standard such as "how good the product performance **should** be." Such a standard has been suggested as an alternative to the expectation standard by various researchers (Morris 1977; Sirgy 1984; Swan, Trawick, and Carroll 1980). The basis for forming the norm level can be varied. It might be based on the average performance level of products in the product category to which the focal product belongs (i.e., product norm). Alternatively it might be based on the performance level of the best brand in that product category (i.e., best-brand norm). Cadotte, Woodruff, and Jenkins (1987) empirically tested these possibilities. Their results suggest that the norm appeared to be based on the average performance level of the product category.

Equity. Consumers purchase a product with an expectation that they will receive at least as equitable a value from the product as what they pay for it. This suggests that during satisfaction formation, consumers may compare input/output combinations in terms of fairness. Thus, in inequitable situations, consumers are likely to express their dissatisfaction (Fisk and Coney 1982; Mowen and Grove 1983; Oliver and DeSarbo 1988). Equity refers to such a normative level of product performance, given the cost they paid (e.g., price).

Ideal. Another alternative standard considered in the literature is ideal or desired. Westbrook and Reilly (1983) argue that consumers are likely to evaluate product performance based on how well they perceive the focal product fulfills their needs and wants. In this respect, the ideal standard represents the optimal product performance a consumer ideally would hope for (Tse and Wilton 1988) or what is personally desired from a product by a consumer (Spreng et al. 1996). Tse and Wilton (1988) empirically demonstrated that ideal was a viable comparison standard during

satisfaction formation.

Comparison of Alternative Standards

Given several types of comparison standards identified, it is an important research issue to understand which type of standard operates best. Or, it might be an even more important issue to identify conditions under which a particular type of standard operates. Up to now only a few studies have empirically compared various types of standards (Cadotte et al. 1987; Tse and Wilton 1988). However, none of them explicitly considered boundary conditions under which a particular standard might be superior to others in explaining satisfaction formation processes.

For example, Cadotte et al. (1987) considered three alternative standards: brand expectation, product norm, and best-brand norm. Comparing three standards in terms of the power to explain the variance in satisfaction with restaurants, they found the product norm superior to the other standards in two out of three situations. On the other hand, Tse and Wilton (1988) tested the relative power of expectation, ideal, and equity in explaining satisfaction in the context of miniature record players. Their results seem to indicate that the brand expectation was better than the ideal or the equity.

Certainly, results from these studies indicate that different standards operate in different consumption situations (i.e., restaurants versus record players). However, it is hard to infer from the studies what the boundary conditions are under which various standards are localized. For alternative standards were not simultaneously compared across a variety of settings (Cadotte et al. 1987). Furthermore, some methodological differences make it difficult to directly compare results between the studies. First, as noted above, the types of alternative standards considered were different across the studies. Second, causal models of satisfaction formation processes adopted in the studies were also different. For example, Tse and Wilton (1988) used a "full-path" model which allowed both direct and indirect paths from the standard to satisfaction and from the perceived performance to satisfaction. On the other hand, Cadotte et al.'s (1987) model constrained the paths such that the comparison standard and the product

performance were to influence satisfaction only indirectly through the subjective disconfirmation. Third, while Cadotte et al. (1987) used different subjective disconfirmations corresponding to different standards in alternative causal models (i.e., disconfirmation from expectation, disconfirmation from product norm, disconfirmation from best-brand norm), Tse and Wilton used only one type of subjective disconfirmation.

The studies to be reported were designed to overcome some of the limitations of previous research in speculating which variables potentially moderate the validity of alternative standards. First, four comparison standards (expectation, product norm, ideal, and equity) were simultaneously tested in four different settings. Second, the same full-path causal model (differing only in the type of comparison standard) was employed to test alternative standards across all settings. The choice of a full-path model was based on previous studies suggesting that a comparison standard and perceived product performance would influence satisfaction levels not only indirectly through disconfirmations but also directly at least in some situations (Bearden and Teel 1983; Bolting and Woodruff 1988; Bolton and Drew 1991; Churchill and Surprenant 1982; Oliver 1980; Oliver and DeSarbo 1988; Swan and Trawick 1981; Tse and Wilton 1988). Third, four types of subjective disconfirmations were measured corresponding to four different comparison standards. Fourth, all the constructs were measured by two response scales and the reliability of each construct was assessed. Finally, all the measures were deliberately kept virtually identical across all settings.

Another important issue to be addressed in this research is a possibility that consumers may use multiple comparison standards simultaneously to evaluate product performance. This possibility has been previously observed by Tse and Wilton (1988). Their results indicated that incorporating a multiple-standard causal-path model (where expectation, ideal, and equity were simultaneously incorporated into a causal-path model) explained more variance in the data than did any of the single-standard models (Spreng et al. 1996). Our research attempted to further test such possibility.

METHOD

Overview of Data Collection

Four satisfaction studies. The data reported here were collected through four consumer satisfaction surveys sponsored by a national food company in Korea. The company is manufacturing and marketing a variety of general food items (e.g., tofu and noodles) and health supplementary food items (e.g., calcium and aloes). Most of them are positioned as high-quality/high-price products, and targeted mainly to those with high education, high income, and health concern. In the first survey, consumer satisfaction levels and their antecedents were measured with respect to the company's tofu, its major product line. The second and third surveys were about the noodle and calcium products. The fourth survey dealt with dealer satisfaction: store managers' satisfaction with the company's general food products was measured. Four different surveys were intended to provide various settings in which models of alternative standards were compared in terms of power to explain the consumer satisfaction process.

Subjects. Subjects for the surveys about tofu and noodles were housewives living in a metropolitan city in Korea. Sampled were only those who had purchased the focal products at least once during the last four weeks before the time of the surveys. A quota sampling procedure was used for sampling based on the demographic profiles of the population such as age and residential area. The sample sizes were 600 and 260 housewives for tofu and noodle products, respectively. For calcium products which were sold mainly by the company's sales force, a customer list was available. Thus, a probability sampling procedure was used for selecting subjects for the survey. In total, 300 subjects were randomly sampled. Finally, a list of supermarkets and stores which carried the company's general food items was used for sampling dealer subjects. 300 store managers were selected through a quota sampling procedure.

Survey Procedure. The surveys were administered by a professional survey organization. The data were collected through a

face-to-face interview with each subject assisted by a structured questionnaire. All interviewers were female, well-trained, and received a detailed orientation about the purpose of the surveys and the contents of the questionnaires. Surveys for tofu and noodle products were conducted first and took 14 days. A month later, surveys for calcium products were conducted. 12 days were taken for completion. Finally, the store managers were surveyed for 16 days.

Survey Instrument

Key Constructs. For the purpose of this research, a number of constructs were necessary to be measured. First, overall satisfaction levels with the focal product and actual product performance levels experienced by subjects were measured. Second, to comparatively test alternative comparison standards against one another, it was necessary to measure four types of comparison standards (expectation, product norm, ideal, and equity) simultaneously from each subject. Third, subjective disconfirmations corresponding to these standards were measured separately. In addition, several background variables were measured including consumer involvement levels and product experience. Two response scales were used to measure each construct (except product experience). They were of 9-point bipolar semantic differential scales and selected based on a review of previous studies and on several pretests during questionnaire construction. These measures are explained below.

Measures. First, overall satisfaction of the product was measured by two scales: ('1' - dissatisfied very much, '9' - satisfied very much) and ('1' - very bad purchase, '9' - very good purchase). After this, the perceived performance level of the product was measured. Subjects were asked about the product quality they actually experienced during consumption regardless of their satisfaction/dissatisfaction level. Two response scales were used: ('1' - very bad product, '9' - very good product) and ('1' - poor quality, '9' - excellent quality').

Next, four alternative comparison standards and corresponding disconfirmation levels were measured. Subjects indicated the level of each

comparison standard on two measurement scales. Then, they evaluated the product performance experienced against that standard level. Thus, the level of each comparison standard and that of corresponding disconfirmation were measured twice. These are explained in a more detail below.

Subjects were first asked about the level of product quality which they had expected prior to purchase would be (expectation). Two scales were provided: ('1' - very bad, '9' - very good) and ('1' - poor quality, '9' - excellent quality). Immediately after each response, subjects evaluated the experienced quality compared to that expectation level (expectation-disconfirmation) using the scale, ('1' - much worse, '9' - much better). After this, the average quality of the product category subjects assumed was measured (product norm) using two response scales: ('1' - very bad, '9' - very good) and ('1' - poor quality, '9' - excellent quality). This was followed by evaluations of the experienced quality against that norm level (norm-disconfirmation). Next were measures of the product quality which consumers personally would desire to receive (ideal). Two response scales were provided: ('1' - average level, '9' - world best level) and ('1' - fair quality, '9' - superb quality). This was followed by evaluations of the experienced quality against that ideal level (ideal-disconfirmation). Finally, subjects reported the product quality which they assumed should be in light of the price they had paid (equity) on two scales: ('1' - very bad, '9' - very good) and ('1' - poor quality, '9' - excellent quality). Again, this was followed by evaluations of the experienced quality against that equity level (equity-disconfirmation).

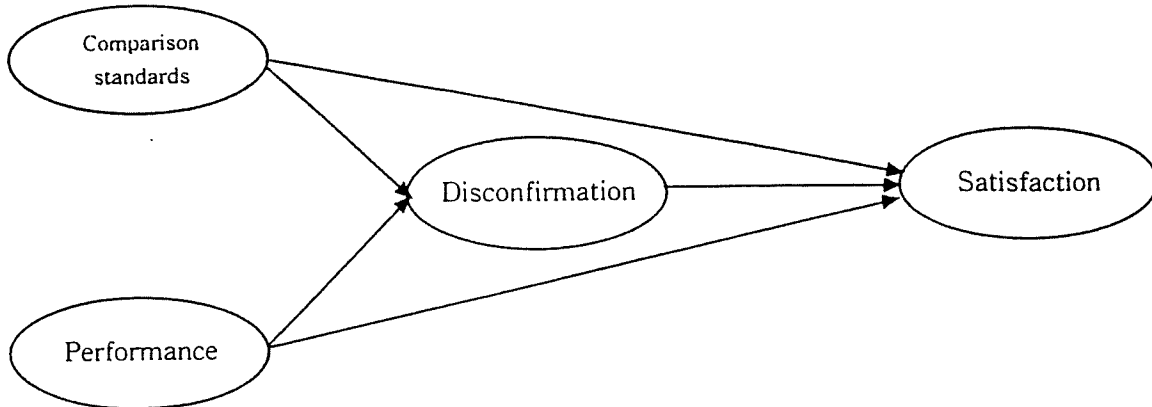
Finally, two response scales were used in order to measure consumer involvement levels with the product: ('1' - not at all interested in the product, '9' - very much interested in the product) and ('1' - not at all careful in choosing a brand, '9' - very much careful in choosing a brand). To measure product experience, the average frequency of purchase per week and the average amount of each purchase were asked by open-ended questions.

Causal Model and Validation Test

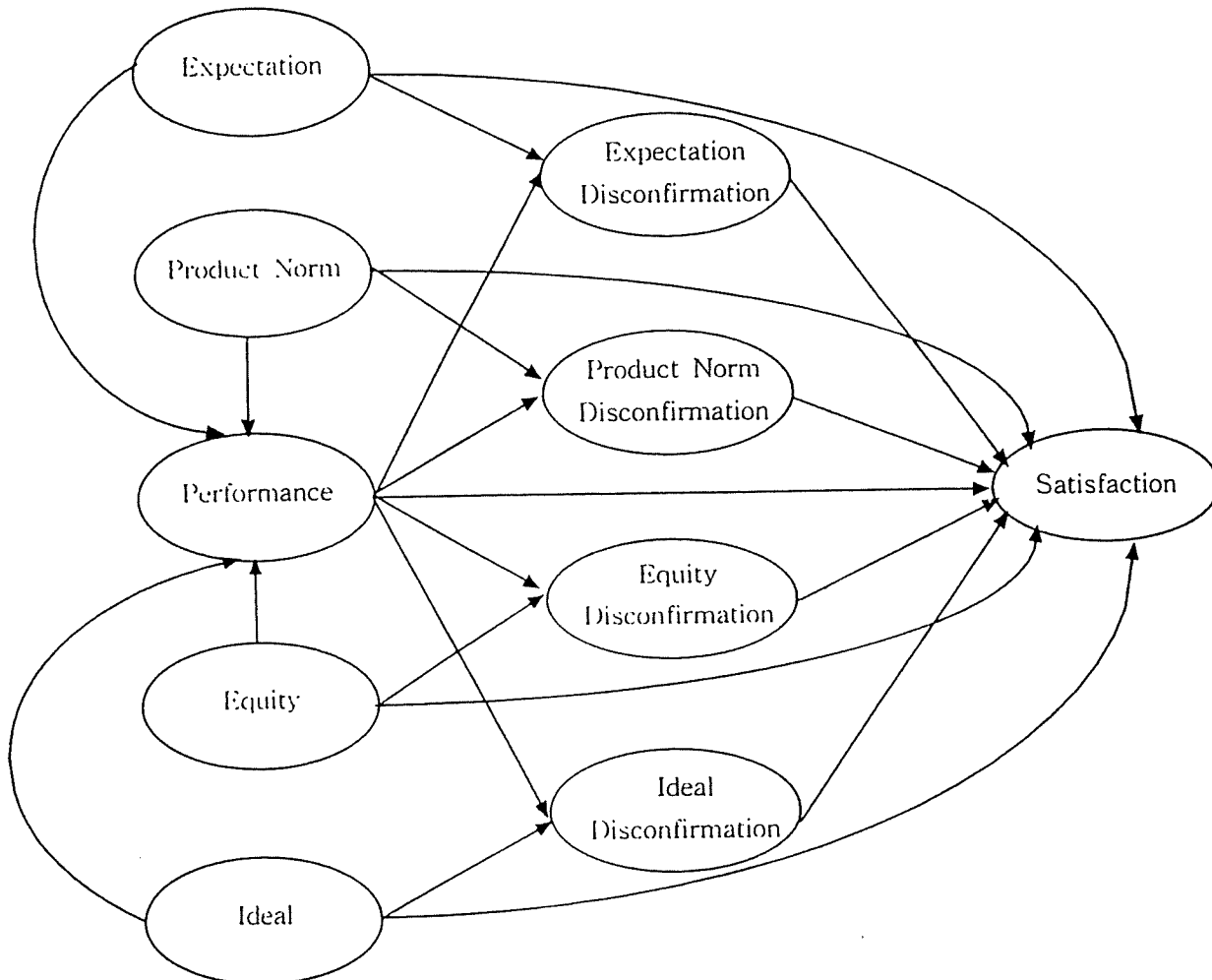
A full-path model was employed to develop

Figure 1
Causal Models of Satisfaction Formation Process

(a) single-standard model



(b) multiple-standard model



alternative causal models of the satisfaction processes. In total, five alternative models were created. First, four single-standard causal models were developed. Figure 1a shows hypothetical paths among the constructs for these models. Actual models tested differed only in the type of standard employed (expectation, norm, ideal, or equity). Next, a multiple-standard model was created in light of a possibility that consumers might use four standards simultaneously to evaluate product performance (Figure 1b). Then, the validity of the five alternative models was assessed using LISREL 8 (Joreskog and Sorbom 1993). To compare the models' validity, five conventional model-fit indicators were used: the chi-square statistic, the goodness-of-fit index (GFI), the adjusted goodness-of-fit index (AGFI), the root mean square residual (RMR), and normed fit index (NFI).

RESULT

Results to be reported here are organized as following. First, results from the causal modeling analyses about the tofu data are presented. Then, potential variables moderating the type of standard used will be speculated. After this, results about the other data (for the noodle, calcium and the dealer satisfaction) will be reported and integrated.

Tofu Data

Reliability and validity of measures. As explained previously, all constructs included in the causal model of consumer satisfaction formation processes were measured using two scale items. The scores obtained from two items for each construct were averaged into a composite score for further analyses. First, reliability was assessed by Cronbach's coefficient alpha for each construct. Analyses revealed that all the measures were highly reliable (all alpha values were higher than .86). Second, a factor analysis was performed on eight response scales measuring four comparison standards (two per each standard), in order to see if the four standards considered were really distinct constructs. A principal component factor analysis with varimax rotation revealed that four factors emerged and they accounted for 90.1 percent of the total variance of the original items. This

supports the validity of the measures.

Table 1
Factor Loadings for Four-Factor Solution

Items	Factor 1	Factor 2	Factor 3	Factor 4
equity 1*	.938			
equity 2*	.938			
norm 1		.951		
norm 2		.946		
expectation 1			.908	
expectation 2			.899	
ideal 1				.916
ideal 2				.873

Note. Items 1 and 2 represent two scale items measuring each standard.

Table 2
Model-Fit Indicators for Alternative-Standard Models in Tofu Data

Indicator	Comparison standard				
	Exp	Norm	Ideal	Equity	All*
Chi-square	42.61	35.49	17.47	50.13	907.10
(p value)	(.000)	(.001)	(.23)	(.000)	(.000)
GFI	.98	.98	.99	.98	.84
AGFI	.95	.96	.98	.94	.77
RMR	.018	.016	.011	.016	.07
NFI	.99	.99	1.00	.99	.91

* Multiple-standard model

Overall model tests. Five causal models of satisfaction formation processes were tested using the LISREL procedure. Table 2 shows the values of various indicators conventionally used for comparing alternative causal models. Two notable patterns emerged. First, all the indicators suggest that the multiple-standard model is inferior to any of the single-standard models in explaining the variance of the data. The multiple-standard model has the highest significant chi-square value, indicating the lowest model fit. Second, results favors the ideal-standard model most among four single-standard models. Chi-square statistics show that only the ideal-standard model had an insignificant value. Further, the model had the

lowest RMR as well as the highest GFI, AGFI, and NFI, representing a very good fit. In sum, causal modeling analyses supported the ideal as the best standard in the tofu case.

Subgroup analysis. Although the analysis of the tofu data strongly supported the ideal-standard model, it was still possible that there were subgroups of subjects who might have employed a different standard to evaluate product performance. We attempted to explore this possibility by dividing subjects into four groups based on the levels of involvement and product experience and testing alternative causal models for each subgroup. The sample size of the data ($n=600$) was deemed sufficiently large to allow this analysis. First, a composite involvement score was calculated by averaging responses on two scale items (Cronbach alpha = .85). Second, a product experience score was obtained by multiplying the average frequency of purchase per week by the average amount of each purchase. Using a median split, four subgroups of subjects were formed: low-experience/low involvement group, high-experience/low-involvement group, low-experience/high-involvement group, and high-experience/high-involvement group. Then, five alternative causal models were tested for each group in turn. Summary results are shown in Table 3.

Some notable patterns emerged in the chi-square statistics. First, the multiple-standard model did not fit the data, thus failing to explain the consumer satisfaction processes. This parallels the results from the overall analysis. Second, the ideal-standard model fit the data in three out of four subgroups, representing the main effect for the tofu product's characteristics. Third, and importantly, the standards supported other than the ideal varied depending on the involvement and experience levels. Specifically, in the high-experience/low-involvement group (C), the expectation-standard model was strongly supported and even better than the ideal-standard model. In the high-experience/high-involvement group (D), two normative-standard models were also supported. The norm- and the equity-standard model fit the data pretty well. By contrast, in the low-experience/high-involvement group (B) the data supported the ideal-standard model only.

Finally, none of the models was statistically supported when both involvement and experience were low (A), although the ideal-standard model was relatively superior to others.

Summary. The overall analysis of the tofu data was supportive of the ideal-standard model. However, the subgroup analyses appeared to suggest that involvement and product experience might interactively influence the type of standard used during satisfaction formation. Briefly, expectation might become a dominant standard when the level of product experience is high but the involvement level is low. On the other hand, normative standards such as product norm and equity might operate well if the levels of both product experience and involvement are sufficiently high. Finally, the ideal standard would best explain the consumer satisfaction processes when involvement is high but experience is low.

As such, our results might suggest interactive effects of involvement and product experience on consumer satisfaction processes. Although this interpretation might be in line with moderating roles of involvement and experience typically found in studies concerning other domains of consumer information processing (Petty, Unnava, and Strathman 1991), it is inevitably speculative in nature. Our interpretation might be bolstered to some extent, however, if similar results are obtained in different situations of consumer satisfaction. For this reason, we attempted to test four single-standard models with the data from the surveys for the noodle and calcium products, as well as for the dealer satisfaction.

Noodle, Calcium, and Dealer Satisfaction Data: Subgroup Analysis

The procedure used in the subgroup analyses was virtually identical to that of the tofu data: (1) dividing subjects into subgroups of low-experience/low involvement, high-experience/low-involvement, low-experience/high-involvement, and high-experience/high-involvement and (2) testing four alternative causal models for each subgroup. It might be ideal to perform this analysis for each of the noodle, calcium, and dealer data sets separately. Unfortunately, the total sample size of each data set was relatively small (n

Table 3
Model-Fit Indicators for Alternative-Standard Models in Tofu Subgroups

Indicator	(a) Low experience and Low involvement (n* = 174)					(b) Low experience and High involvement (n = 134)				
	<u>Exp</u>	<u>Norm</u>	<u>Ideal</u>	<u>Equity</u>	<u>All</u>	<u>Exp</u>	<u>Norm</u>	<u>Ideal</u>	<u>Equity</u>	<u>All</u>
Chi-square	52.60	43.68	26.55	49.79	516.71	33.49	21.98	18.62	38.90	361.18
(p value)	(.000)	(.000)	(.022)	(.000)	(.000)	(.002)	(.079)	(.18)	(.000)	(.000)
GFI	.93	.95	.96	.93	.77	.94	.96	.97	.94	.79
AGFI	.83	.86	.90	.83	.66	.85	.90	.92	.84	.70
RMR	.037	.034	.024	.03	.078	.031	.016	.026	.024	.051
NFI	.95	.96	.97	.95	.84	.97	.98	.98	.97	.89

Indicator	(c) High experience and Low involvement (n = 121)					(d) High experience and High involvement (n = 127)				
	<u>Exp</u>	<u>Norm</u>	<u>Ideal</u>	<u>Equity</u>	<u>All</u>	<u>Exp</u>	<u>Norm</u>	<u>Ideal</u>	<u>Equity</u>	<u>All</u>
Chi-square	13.62	21.84	14.94	46.74	342.03	22.83	18.83	14.37	19.62	348.90
(p value)	(.48)	(.082)	(.38)	(.000)	(.000)	(.063)	(.17)	(.42)	(.14)	(.000)
GFI	.97	.96	.97	.92	.79	.96	.96	.97	.96	.80
AGFI	.93	.89	.92	.79	.69	.89	.90	.93	.91	.70
RMR	.024	.053	.042	.033	.12	.047	.016	.026	.072	.12
NFI	.98	.96	.98	.91	.82	.97	.98	.98	.97	.83

* n represents sample size.

= 300 or less) compared to the sample size of the tofu data (n = 600). Consequently, it might be impractical to test causal models using LISREL for each subgroup because of its limited sample size (n = 75 or less). Our strategy therefore was to combine the three data sets for subgroup analyses. First, for each data set four subgroups were created based on involvement and experience scores. To make subgroups as dissimilar as possible in terms of involvement and experience levels, subjects around the median level of involvement and experience were excluded. Next, the subgroups of the same involvement/ experience condition were merged across three data sets, resulting in four distinct involvement/experience subgroups for the combined data. Finally, four alternative single-standard models were tested for each of these subgroups. If the results from these analyses were consistent with those from the tofu subgroup analyses, then our earlier interpretation regarding interactively effects of involvement and

experience could be bolstered. In fact, this was the case.

Table 4 contains summary results of the subgroup analyses. Consistent with the tofu case, involvement and experience appeared to interactively influence consumer satisfaction processes. First, in the high-experience/low-involvement group (C), the expectation-standard model fit the data best and it was the only one that nearly reached statistical significance. This parallels the result from the tofu data. In the high-experience/high-involvement group (D), however, the norm-standard model performed best in explaining consumer satisfaction processes. This is also consistent with the implications of the tofu subgroup analysis. By contrast, the data predominantly supported the ideal-standard model in the low-experience/high-involvement group (B), whereas none of the models was supported (although the ideal model was most favored) in low-involvement/low-experience group (A). Again,

Table 4
Model-Fit Indicators for Alternative-Standard Models in Subgroups of
Combined Samples from Noodle, Calcium, and Dealer Data Sets

Indicator	(a) Low experience and Low involvement (n [*] =134)				(b) Low experience and High involvement (n=128)			
	Exp	Norm	Ideal	Equity	Exp	Norm	Ideal	Equity
Chi-square	35.65	33.70	26.93	34.82	23.18	58.07	13.41	25.02
(p value)	(.001)	(.002)	(.02)	(.001)	(.057)	(.000)	(.49)	(.034)
GFI	.94	.94	.96	.94	.96	.91	.97	.95
AGFI	.84	.85	.89	.85	.89	.78	.93	.88
RMR	.03	.03	.035	.034	.017	.085	.017	.034
NFI	.96	.96	.97	.95	.97	.93	.98	.97

Indicator	(c) High experience and Low involvement (n=104)				(d) High experience and High involvement (n=118)			
	Exp	Norm	Ideal	Equity	Exp	Norm	Ideal	Equity
Chi-square	24.25	32.88	53.33	32.50	10.02	8.88	9.86	19.15
(p value)	(.043)	(.003)	(.000)	(.003)	(.76)	(.90)	(.77)	(.16)
GFI	.94	.93	.89	.92	.98	.98	.98	.96
AGFI	.84	.81	.71	.78	.95	.96	.95	.90
RMR	.034	.031	.036	.029	.014	.008	.024	.021
NFI	.96	.95	.93	.94	.99	.99	.99	.98

* n represents sample size.

this was also consistent with the results from the tofu subgroup analysis.

DISCUSSION

The extant literature suggests that consumers may use one or some of different comparison standards to evaluate actual product performance during consumer satisfaction formation. However, boundary conditions under which a particular standard operates have not been identified. This study explored such conditions by analyzing data over various consumption situations. Results suggested that (1) consumers might use a single standard rather than multiple standards during satisfaction formation and (2) consumer involvement and product experience might interactively influence the type of comparison standard used in the satisfaction formation process. The summary and implications of the results are

now discussed.

The results suggest that in high-involvement/high-experience situations, a normative standard like product norm is likely to operate. This was initially suggested by the subgroup analysis of the tofu data, and supported by the subgroup analysis of the combined data for the noodle, calcium, and dealer satisfaction. Accordingly, highly involved consumers with high experience might judge the product quality in reference to what they assume the product should provide. This finding seems consistent with Woodruff, Cadotte, and Jenkins (1983). They argued that norms were constrained by the performance consumers believe was possible as indicated by the performance of "known" brands. Consequently, having some experience with the product is a necessary condition to possess a norm standard in memory. On the other hand, a normative standard is related to consumers'

emotional commitment in meeting their needs and wants by purchasing a product. The involvement concept seems to reflect such emotional commitment. A norm standard is then unlikely to be used by uninvolved consumers. Therefore, a normative standard is likely to operate well for highly involved consumers with ample product experience.

The subgroup analysis of the tofu data suggested that the expectation might be a dominant comparison standard to evaluate product performance in low-involvement/high-experience situations. This tended to be further supported by the subgroup analysis of the combined data. This finding appears consistent with some of previous results from information processing research. Specifically, uninvolved consumers tend to simply use easily accessible information in memory to make a judgment (Park and Hastak 1994; Sanbonmatsu and Fazio 1990). Also, memory information such as a prior brand evaluation tends to be more accessible when it is experience-based than when it is information-based (Berger and Mitchell 1989; Fazio and Zanna 1981). Since the expectation of a brand is in fact a previously-formed brand evaluation, it is likely to be easily accessible when the experience of the brand is accumulated. Consequently, relatively uninvolved consumers with high product experience are likely to simply retrieve and use the expectation as a standard to evaluate product performance.

Results also indicated that in high-involvement/low-experience situations the ideal-standard model stood out. This was supported by the subgroup analyses of the tofu and by the analysis of the combined data. This finding seems intuitively reasonable. Highly involved consumers with low experience would be highly concerned with meeting their needs and wants. However, they are unlikely to have a strong normative standard to evaluate product performance due to their lack of product experience. Therefore, they are likely to simply evaluate the product performance against what they personally desire to receive from the product.

On the other hand, no causal model was supported in low-involvement/low-experience situations in the subgroup analysis of either the tofu or the combined data. One possibility is that satisfaction is predominantly determined by the

actual product performance (Cronin and Taylor 1992). However, this clearly needs to be examined by future research.

The final issue to be addressed is the validity of multiple-standard model for consumer satisfaction processes. Previous research has suggested a possibility that consumers might utilize multiple standards to evaluate product performance (Tse and Wilton 1988). In our research, however, the multiple-standard model did not adequately account for the satisfaction formation processes for any of the subgroups. This clearly questions the viability of the multiple-standard model. However, all of the four standards were contained in the multiple-standard model in this research. In this respect, our results cannot rule out a possibility that consumers might use only a subset (not all) of the standards simultaneously. Empirical assessment of such possibility would be enormously taxing as the number of comparison standards considered increases, unless some theory-based predictions are delineated about a particular set of standards operating in a certain situation. Definitely, future research is needed in this direction.

One caveat to our conclusions should be emphasized. Our findings are exploratory in nature. None of the factors considered as moderators were not experimentally manipulated. Specifically, high versus low levels of involvement and product experience variables were operationalized by a median split of measured scores. This certainly introduces a danger of confounding and thus weakens the validity of our interpretations of the results. A replication of our findings is needed. Future research involving experimental manipulations of the variables such as involvement and product experience is warranted.

REFERENCES

- Bearden, William O. and Jesse E. Teel (1983), "Selected Determinants of Consumer Satisfaction and Complaint Reports," *Journal of Marketing Research*, 20, (February), 21-8.
- Berger, Ida E and Andrew A. Mitchell (1989), "The Effect of Advertising on Attitude Accessibility, Attitude Confidence, and the Attitude-Behavior Relationship," *Journal of Consumer Research*, 16, (December), 269-279.
- Bolting, Claire P. and Robert B. Woodruff (1988), "Effects of Situational Involvement on Consumers' Use of Standard in Satisfaction/Dissatisfaction Process,"

- Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 1, 16-24.
- Bolton, Ruth N. and James H. Drew (1991), "A Multistage Model of Customers' Assessments of Service Quality and Value," *Journal of Consumer Research*, 17, (March), 375-84.
- Cadotte, Ernest R., R. B. Woodruff and R. L. Jenkins (1987), "Expectation and Norms in Models of Consumer Satisfaction," *Journal of Marketing Research*, 24, (August), 305-14.
- Churchill, Gilbert A., Jr. and C. Surprenant (1982), "An Investigation into the Determinants of Customer Satisfaction," *Journal of Marketing Research*, 19, (November), 491-504.
- Cronin, J. Joseph, Jr. and Steven A. Taylor (1992), "Measuring Service Quality: A Reexamination and Extension," *Journal of Marketing*, 56, (July), 55-68.
- Day, Ralph L. (1984), "Modeling Choices Among Alternative Responses to Dissatisfaction," *Advances in Consumer Research*, 11, Thomas C. Kinnear, (ed.), Ann Arbor, MI: Association for Consumer Research, 469-79.
- Fazio, Russell H. and Mark P. Zanna (1981), "Direct Experience and Attitude-Behavior Consistency," *Advances in Experimental Social Psychology*, 14, Leonard Berkowitz, (ed.), New York: Academic Press, 161-198.
- Fisk, Raymond P. and Kenneth A. Coney (1982), "Postchoice Evaluation: An Equity Theory Analysis of Consumer Satisfaction/Dissatisfaction with Service Choices," *Conceptual and Empirical Contributions to Consumer Satisfaction Complaining Behavior*, H. Keith Hunt and Ralph L. Day, (eds.), Bloomington, IN: Indiana University School of Business, 9-16.
- Joreskog, Karl and Dag Sorbom (1993), *LISREL 8: A Guide to the Program and Applications*, Chicago: Scientific Software International.
- Morris, Earl W. (1977), "A Normative Deficit Approach to Consumer Satisfaction," *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*, H. Keith Hunt, (ed.), Cambridge, MA: Marketing Science Institute, 240-74.
- Mowen, John C. and Stephen J. Grove (1983), "Search Behavior, Price Paid, and the 'Comparison Other': An Equity Theory Analysis of Post Purchase Satisfaction," *International Fare in Consumer Satisfaction and Complaining Behavior*, Ralph L. Day and H. Keith Hunt, (eds.), Bloomington, IN: Indiana University School of Business, 57-63.
- Oliver, Richard L. (1980), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, 17, (September), 46-9.
- Oliver, Richard L. and G. Linda (1981), "Effects of Satisfaction and Its Antecedents on Consumer Preference and Intention," *Advances in Consumer Research*, Kent B. Monroe, (ed.), Ann Arbor, MI: Association for Consumer Research, 88-93.
- Oliver, Richard L. and Wayne S. DeSarbo (1988), "Response Determinants in Satisfaction Judgements," *Journal of Consumer Research*, 14, (March), 495-507.
- Park, Jong-won and Manoj Hastak (1994), "Memory-based Product Judgement: Effects of Involvement at Encoding and Retrieval," *Journal of Consumer Research*, 21, (December), 534-547.
- Petty, Richard E., Rao Unnava, and Alan J. Strathman (1991), *Handbook of Consumer Behavior*, Thomas S. Robertson and Harold H. Kassarian, (eds.), Englewood Cliffs, NJ: Prentice-Hall, 1-49.
- Sanbonmatsu, David M. and Russell H. Fazio (1990), "The Role of Attitudes in Memory-Based Decision Making," *Journal of Personality and Social Psychology*, 59, (4), 614-622.
- Sirgy, M. Joseph (1984), "A Social Cognition Model of CS/D: An Experiment," *Psychology and Marketing*, 1, (Summer), 27-44.
- Spreng, Richard A., Scott B. Mackenzie and Richard W. Olshavsky (1996), "A Reexamination of the Determinants of Consumer Satisfaction," *Journal of Marketing*, 60, (March), 15-32.
- Swan, John E. and I. Fredrick Trawick (1981), "Disconfirmation of Expectations and Satisfaction with a Retail Service," *Journal of Retailing*, 57, (Fall), 49-67.
- Swan, John E., I. Fredrick Trawick and Maxwell G. Carroll (1980), "Satisfaction Related to Predictive, Desired Expectations: A Field Study," *New Findings on Consumer Satisfaction and Complaining Behavior*, H. Keith Hunt and Ralph L. Day, (eds.), Bloomington, IN: Indiana University, 15-22.
- Tse, David K. and Peter C. Wilton (1988), "Models of Consumer Satisfaction: An Extension," *Journal of Marketing Research*, 25, (May), 204-12.
- Westbrook, Robert A. (1980), "Product/Consumption-Based Affective Responses and Postpurchase Processes," *Journal of Marketing Research*, 24, (August), 258-70.
- Westbrook, Robert A. and Michael D. Reilly (1983), "Value-Percept Disparity: An Alternative to the Disconfirmation of Expectations Theory of Consumer Satisfaction," *Advances in Consumer Research*, Richard P. Bagozzi and Alice M. Tybout, (eds.), Ann Arbor, MI: Association for Consumer Research, 256-61.
- Woodruff, Robert B., E. R. Cadotte and R. L. Jenkins (1983), "Modeling Consumer Satisfaction Processes Using Experience-Based Norms," *Journal of Marketing Research*, 20, (August), 296-304.
- Yi, Youjae (1990), "A Critical Review of Consumer Satisfaction," *Review of Marketing*, Valarie A. Zeithaml, (ed.), Chicago: American Marketing Association, 68-123.

ACKNOWLEDGEMENTS

A part of this article has been presented at the Association for Consumer Research Conference at Denver, Colorado in 1997. The authors would like to thank members of B.E.S.T. of the Korea University for their valuable suggestions concerning the theoretical basis for the study and the interpretation of the results. Special thanks are given to Pulmuone Co. for supporting the data collection and to the editor for his valuable guidance.

Send correspondence regarding this article to:

Jong-Won Park
Department of Business Administration
Korea University
Seoul, KOREA

THE INFLUENCE OF AFFECTIVE STATE ON SATISFACTION RATINGS

Beverly K. Brockman, University of Alabama

ABSTRACT

The author examines the influence of affective state – both temporary affect unrelated to the product/service being evaluated and life satisfaction – on consumer satisfaction ratings. Temporary affect is manipulated and then separated into two states, positive and negative. Three different scales are used to measure consumer satisfaction with the most recent shoe purchase and restaurant experience: 1) an adaptation of Oliver's (1997) consumption satisfaction scale, 2) the single-item delighted/terrible scale developed by Andrews and Withey (1976) and 3) a single-item, bipolar, very satisfied/very dissatisfied scale. The results indicate that the Oliver (1997) scale is the most robust against influence from temporary positive affect and life satisfaction. No influence from negative affect was found.

INTRODUCTION

The past two decades have seen a significant amount of empirical study intended to further our understanding of consumer satisfaction. These research efforts resulted in the predominance of the confirmation/disconfirmation paradigm (Yi 1990), which assumes satisfaction arises from cognitive processes of comparison. Despite our focus on cognitive processes, the role of affect in satisfaction has also been considered for quite some time (Isen, Shalke, Clark, and Karp 1978; Westbrook 1980; Westbrook 1987; Westbrook and Oliver 1991; Mano and Oliver 1993). A review of the various definitions of consumer satisfaction (Yi 1990) reveals disagreement with regard to the role of affect within the construct. For example, Howard and Sheth (1969, p. 145) define consumer satisfaction as "the buyer's *cognitive* state of being adequately or inadequately rewarded for the sacrifices he has undergone" (emphasis added). Thus, satisfaction is defined as purely cognitive. In contrast, Oliver (1981, p.27) defines consumer satisfaction as "the summary psychological state resulting when the emotion surrounding disconfirmed expectations is coupled with the

consumer's prior feeling about the consumption experience." Oliver clearly includes affect in his definition. In his most recent work, Oliver (1997, p.13) presents a revision of his earlier definition. "*Satisfaction* is the consumer's fulfillment response. It is a judgment that a product or service feature, or the product or service itself, provided (or is providing) a *pleasurable* level of consumption-related fulfillment, including levels of under- or overfulfillment" (emphasis in original). His latest definition still incorporates affect, but the word *judgment* implies a cognitive process as well. In fact, Oliver (1997) actually presents different forms of satisfaction, each of which incorporates different types of affect and cognitive processes.

In order to adequately define consumer satisfaction we must understand the role affect plays within it. Another important question, however, involves the influence of unrelated affect on the *measurement* of satisfaction. In other words, how much does one's affective state at the time of reporting influence his or her satisfaction rating? Are certain measures more susceptible to the influence of affective state unrelated to the product/service being evaluated? It is these questions that the author has addressed in this study. We are aware that affect may arise from the consumption experience (Westbrook 1987; Westbrook and Oliver 1991); we are also aware that it may act as a dimension in the formation of satisfaction (Alford and Sherrell 1996). This study, however, considers unrelated affect experienced at the time of reporting as a nuisance variable, which debilitates the *measurement* of satisfaction. Temporary affect is manipulated and then separated into two states, positive and negative. Thus, the influence of both states on the satisfaction rating is assessed.

The role of one's general disposition, or life satisfaction, has also been considered in consumer satisfaction research (Westbrook 1980; Peterson and Wilson 1992). Most work has focused on the correlation between life satisfaction and the satisfaction rating. While this is helpful, the researcher cannot separate the different roles life satisfaction might play in the true satisfaction

evaluation and the satisfaction rating, a measurement issue. This author did not distinguish between these two different roles either, but the manipulation and measurement of temporary affect enables us to examine the interaction between the two variables: life satisfaction and temporary affective state unrelated to the product/service being evaluated. The research also extends previous studies by comparing the susceptibility of several different satisfaction scales to affective state, and by investigating the separate influence of positive and negative affect.

LITERATURE REVIEW

Westbrook (1980) used regression analysis to determine the influence of expectations, level of optimism/pessimism, overall life satisfaction, consumer discontent, and mood on automobile and footwear satisfaction. The results indicated support for relationships between automobile satisfaction and the realization of expectations, overall life satisfaction, and consumer discontent; neither optimism/pessimism nor mood displayed significant relationships. The results for footwear satisfaction only indicated support for a relationship with expectancy realization; none of the hypothesized affective influences appeared to be related. One reason given by Westbrook for the absence of the hypothesized effect of mood on satisfaction was the stable environment in which the research was conducted. The subjects of the study were undergraduate students at the University of Arizona, enrolled in an introductory business course. A questionnaire was administered during class, which included measures of product satisfaction, dispositional sources of affect, and concurrent affective states. Moods extreme enough to influence satisfaction ratings were probably not experienced by the subjects since no manipulation of affective state was made and the subjects completed the questionnaire in a stable situation.

Peterson and Wilson (1992) used a telephone survey to test the relationship between subjective well being, satisfaction with a vehicle, and mood. Correlations between the three variables revealed statistically significant positive relationships between life satisfaction and mood, life satisfaction

and vehicle satisfaction, and vehicle satisfaction and mood. The correlation between vehicle satisfaction and mood (.19) was smaller than the correlation between vehicle satisfaction and life satisfaction (.27). Intuitively, the smaller correlation makes sense because an individual's mood at a certain moment would be expected to have a stronger relationship with their satisfaction with life in general than with their satisfaction with a product that can be replaced. Additionally, the small correlation between vehicle satisfaction and mood could be expected after reviewing the results of Westbrook's (1980) study. Affect manipulation was not conducted in either study, but Peterson and Wilson's respondents did complete the survey in different environments. Thus, the respondents' affective state had enough variance to have a slight influence on satisfaction ratings.

A significant amount of research involving manipulation of affective state can be found in the psychology literature; however, those studies conducted for the explicit purpose of determining the influence of affect on satisfaction ratings are rare. Isen et al. (1978) found that manipulation to induce good moods (a free gift), resulted in higher satisfaction ratings for automobiles and television sets. The reason suggested for these results is that positive feelings cue positive memories, which then influence judgment.

A thorough literature review produced no studies involving manipulation of negative affect to determine its effect on consumer satisfaction ratings. One reason for this could be that in research experiments the influence of manipulated negative affect on behavior has been less significant than that of manipulated positive affect. Isen (1984) discussed the conflicting evidence concerning the impact of negative affect on behavior. Positive affect manipulations have been shown to result in more positive judgment and behavior. Negative affect manipulations sometimes result in thought and behavior opposite of those with positive affect (Cialdini and Kenrick 1976; Moore et al. 1973; Weyant 1978); other times negative affect results in the same kinds of behavior as that produced by positive affect (Isen 1970; Mischel et al. 1968; Mischel et al. 1976). One reason proposed for this discrepancy is that individuals in a negative affective state attempt to improve their mood through positive actions

(Cialdini et al. 1973; Isen, et al. 1973; Weyant 1978). Individuals in a positive mood, however, strive to maintain the nice feeling.

Westbrook (1987), Westbrook and Oliver (1991), and Mano and Oliver (1993) studied the role of both positive and negative affect in satisfaction. Their research, however, was approached from the theory that the consumption experience elicits certain emotions, which then influence the postpurchase product/service evaluation. It is worth noting that Westbrook (1987) acknowledged Isen's (1984) interpretation of the effects of affective state on cognitive retrieval processes, which includes the possibility of affective state at the time of reporting systematically biasing satisfaction ratings. Thus, true satisfaction derived from consumption could interact with the temporary affective state felt at the time the satisfaction measurement was taken.

Any conclusions that can be drawn from prior research regarding the influence of life satisfaction and positive and negative affect on satisfaction ratings are shaky at best. Although two studies (Westbrook 1980; Peterson and Wilson 1992) found an influence of life satisfaction on automobile satisfaction ratings, Westbrook (1980) found no such influence on footwear satisfaction ratings. It seems plausible that life satisfaction may have more influence on ratings of high involvement products and services. At this point, however, such predictions are more speculative than conclusive.

Similar conflict in research results has been found in studies of temporary affect and its influence on satisfaction ratings. Westbrook's (1980) initial study did not reveal any influence of mood state on satisfaction ratings with either automobiles or footwear. An influence of positive affect on satisfaction ratings of automobiles and television sets was found, however, by Isen et al. (1978) when positive affect was manipulated. Peterson and Wilson (1992) also found influence of mood state on satisfaction ratings for automobiles. These authors did not manipulate affect, nor separate its measurement into positive and negative categories. The greater variance in mood found in this study, compared to Westbrook (1980), can be attributed to the respondents' completion of the survey instrument in different environments. Thus, there is some evidence of

influence on satisfaction ratings from temporary mood state in general and positive affect in particular; no influence of negative affect has been found.

Due to the conflict in prior research results, and the lack of research into the influence of negative affect on satisfaction ratings, no formal hypotheses were developed for this study. Rather, it is intended to be a more thorough exploration into the role life satisfaction and affective state play in the measurement of satisfaction. It is hoped that the results of this study lead to a better understanding of these variables in the context of satisfaction measurement.

METHOD

The subjects of the study were undergraduate students enrolled in finance and marketing courses at a major university located in the southeast United States. After excluding unusable surveys the sample size was 228, broken down into 54.7% male and 45.3% female. All subjects first completed the Diener, Emmons, Larsen, and Griffin (1985) satisfaction with life scale. After completing this measure, each class was then subjected to one of three mood manipulations: positive, negative, or control (no manipulation). The purpose of the manipulation was to insure that enough variance with the temporary affective state variable was obtained to determine its influence on satisfaction ratings.

The positive manipulation group was shown an uplifting video and given candy as appreciation for participating in the study. The negative manipulation group was shown a sad video; immediately after the subjects viewed the video, the researcher made a few brief comments about the seriousness of the subject matter - how the families of plane crash victims are treated by the airline. Both positive and negative videos also included a commercial at the beginning. A positive commercial for O'Douls nonalcoholic beer was shown to the positive manipulation group; a serious commercial for medical books published by Time Life Books was shown to the negative manipulation group.

Both positive and negative manipulation groups were told that they were helping with two projects. One concerned satisfaction ratings for

life, products, and services. The other concerned the effectiveness of television advertisements when viewed with certain types of news segments. To make the cover story more believable, the survey instrument included questions about the commercials. The control group was not subjected to a mood manipulation. They were simply told that they were helping with a project regarding satisfaction with life, products, and services.

Once the mood manipulations were conducted, the students completed a survey instrument which consisted of the following scales: 1) the PANAS scale developed by Watson et al. (1988), used to measure positive and negative affect felt at the time the survey was completed, 2) an adaptation of Oliver's (1997) consumption satisfaction scale, measured on a five point scale, 3) the single-item

EXHIBIT 1

<p align="center">Satisfaction With Life Scale (Diener, Emmons, Larsen, and Griffin 1985)</p> <p>7-point scale: 1= strongly disagree, 2= disagree, 3= slightly disagree, 4= neither agree nor disagree, 5= slightly agree, 6= agree, 7= strongly agree</p> <ol style="list-style-type: none"> 1. In most ways my life is close to my ideal. 2. The conditions of my life are excellent. 3. I am satisfied with my life. 4. So far I have gotten the important things I want in life. 5. If I could live my life over, I would change almost nothing.
<p align="center">Shoe and Restaurant Satisfaction (Adapted from Oliver's (1997) consumption satisfaction scale)</p> <p>5-point scale: 1 = strongly agree, 2 = agree, 3 = neither agree nor disagree, 4 = disagree, 5 = strongly disagree</p> <ol style="list-style-type: none"> 1. This is one of the best pair of shoes I could have bought. 2. This pair of shoes is exactly what I need. 3. This pair of shoes hasn't worked as well as I thought they would. 4. I am satisfied with my decision to buy this pair of shoes. 5. Sometimes I have mixed feelings about wearing this pair of shoes. 6. My choice to buy this pair of shoes was a wise one. 7. If I could do it over again, I'd buy a different style/brand. 8. I have truly enjoyed this pair of shoes. 9. I feel bad about my decision to buy this pair of shoes. 10. I am not happy that I bought this pair of shoes. 11. Owning this pair of shoes has been a good experience. 12. I'm sure I made the right choice in buying this pair of shoes. <p>Note: The restaurant measure used the same items with "restaurant" substituted for "shoes"</p>
<p align="center">Shoe and Restaurant Satisfaction (Andrews and Withey 1976)</p> <p>7 point scale: 1 = delighted, 2 = pleased, 3 = mostly satisfied, 4 = mixed about equally satisfied and dissatisfied, 5 = mostly dissatisfied, 6 = unhappy, 7 = terrible</p> <ol style="list-style-type: none"> 1. How do you feel about the purchase of these shoes (restaurant)?
<p align="center">Shoe and Restaurant Satisfaction (Overall Satisfaction Measure)</p> <p>7 point scale: 1 = very satisfied, 4 = satisfied, 7 = very dissatisfied</p> <ol style="list-style-type: none"> 1. Overall, how satisfied have you been with this pair of shoes (restaurant)?
<p align="center">Temporary Positive and Negative Affective State (Watson, Clark, and Tellegen 1988)</p> <p>Respondents indicated how they felt at the present time for 20 different emotions, 10 for positive affect, and 10 for negative affect</p> <p>5 point scale: 1 = very slightly or not at all, 2 = a little, 3 = moderately, 4 = quite a bit, 5 = extremely</p> <p><u>Positive Affect Emotions</u> Interested, Excited, Strong, Enthusiastic, Proud, Alert, Inspired, Determined, Attentive, and Active</p> <p><u>Negative Affect Emotions</u> Distressed, Upset, Guilty, Scared, Hostile, Irritable, Ashamed, Nervous, Jittery, Afraid</p>

“delighted-terrible” scale developed by Andrews and Withey (1976), measured on a seven point scale, and 4) a single-item, seven point, bipolar rating scale (very satisfied/very dissatisfied). The subjects completed the measures in regard to their satisfaction with their most recent shoe purchase and restaurant experience. The measures used in the study are shown in Exhibit 1.

RESULTS

Manipulation Checks

Dissipation of Temporary Affect. The measure of positive and negative affect, the PANAS scale, also served as a manipulation check. Perdue and Summers (1986) discussed solutions to the problems found in using manipulation checks within the main experiment. The counterbalancing approach, in which half of the subjects complete the check before the dependent measures and the other half afterward was one approach suggested by these authors. This method is particularly applicable for this study because of the inconsistencies found with negative affect manipulations conducted in previous research.

A t-test was used to check for differences in negative and positive affect based on the placement of the manipulation check. For the sample as a whole, with combined results of the three manipulation groups, no significant differences in negative and positive affect scores were found between groups who completed the manipulation check at the beginning of the survey versus the end. A t-test conducted within each individual manipulation treatment group, however, revealed an interesting change in negative affect in the negative manipulation treatment group.

The mean negative affect score for those respondents who completed the manipulation check at the beginning of the survey was 19.63; this same score for those respondents who completed the manipulation check at the end of the survey was 15.38. The difference between these two groups was significant at $p = .005$. Thus, the negative affect from the manipulation dissipated within the short period of time that the respondents completed the survey, approximately ten minutes. A significant difference ($p = .032$) was also found

in the control group for the difference in positive affect between the front (mean positive affect score = 28.08) and back (mean positive affect score = 23.23) manipulation check groups. No dissipation of positive or negative affect was found in the positive manipulation treatment group. Two-way Anova's revealed a significant interaction between the manipulation treatment group and the manipulation check placement. The results from these manipulation checks are given in Exhibits 2 through 4.

Difference in Affective State Between the Treatment Manipulation Groups. A one-way ANOVA revealed a significant difference in positive affect between the positive affect manipulation group and the negative affect manipulation group, while there was not a significant difference in negative affect between any of the three treatment groups. The quick dissipation of negative affect appears to have eliminated any significant difference in this affective state between the manipulation treatment groups. The results of this manipulation check are given in Exhibit 5.

Satisfaction Ratings

Correlations were conducted to determine the influence of temporary positive and negative affect and life satisfaction on product and service satisfaction ratings. A significant, positive correlation was found between life satisfaction and shoe satisfaction measured with the Andrews and Withey (1976) delighted/terrible scale. Additionally, a significant, positive correlation was found between positive affect and shoe satisfaction measured with the single item scales, the delighted/terrible measure and the overall very satisfied/very dissatisfied measure. A significant correlation was not found between negative affect and any of the restaurant or shoe satisfaction measures. Thus, significant correlations were only found with shoe satisfaction, and these relationships were only maintained with positive affect and life satisfaction. Additionally, no significant correlations were found with the adapted Oliver (1997) measure. The results of the correlation analysis are given in Exhibit 6.

A one-way ANOVA was conducted to identify

Exhibit 2
Dissipation of Manipulated Affect for Combined Sample T-Tests

Affective State	Manipulation Check Placement	Mean	Significant Difference (p value)
Positive Affect	Front	27.46	.176
	Back	25.99	
Negative Affect	Front	16.26	.332
	Back	15.50	

Exhibit 3
Dissipation of Manipulated Affect Within Each Treatment Group T-Tests

Negative Affect

Manipulation Check Placement	Negative Affect Manipulation Treatment	Positive Affect Manipulation Treatment	Control Group Manipulation Treatment
Front	19.63	15.62	15.14
Back	15.38	15.27	16.18
Significant Difference (p value)	.005	.803	.416

Positive Affect

Manipulation Check Placement	Negative Affect Manipulation Treatment	Positive Affect Manipulation Treatment	Control Group Manipulation Treatment
Front	25.13	28.11	28.08
Back	24.40	29.89	23.23
Significant Difference (p value)	.681	.319	.032

significant differences in satisfaction ratings by manipulation group; however, a test of homogeneity of variances revealed a violation of the equal variance assumption. Therefore, the Kruskal-Wallis non-parametric test was conducted to identify significant differences in satisfaction ratings by manipulation group. Significant differences ($\alpha = .10$) were found between the positive and negative manipulation groups for the overall measure and the delighted/terrible measure of shoe satisfaction. Using these measures, the negative affect manipulation group had a significantly lower satisfaction rating than the positive affect manipulation group. The results of this test are given in Exhibit 7.

To further test the practical significance of the correlations between life satisfaction and positive and negative affect with the shoe and restaurant satisfaction ratings, multiple regression analysis was conducted. Six multiple regressions were run with each shoe and restaurant satisfaction rating as the dependent variable. To eliminate problems with multicollinearity, the remaining measures were combined into a revised shoe satisfaction variable and a revised restaurant satisfaction variable. The combinations varied depending on which measure served as the dependent variable. Temporary affective state did not indicate a significant coefficient in any of the multiple regression analyses. These results are consistent

Exhibit 4

Difference in Positive Affective State: By Manipulation Treatment Group and Manipulation Check Placement Two-Way Anova

		F	Significance
Main Effects	(Combined)	4.423	.005
	Treatment Group	5.671	.004
	Manipulation Check	1.298	.256
2-way Interactions	Treatment Group * Manipulation Check	3.030	.050
Model		3.682	.003

Difference in Negative Affective State: By Manipulation Treatment Group and Manipulation Check Placement Two-Way Anova

		F	Significance
Main Effects	(Combined)	2.074	.105
	Treatment Group	2.645	.073
	Manipulation Check	2.172	.142
2-way Interactions	Treatment Group * Manipulation Check	3.752	.025
Model		2.461	.034

Exhibit 5

Difference in Affective State by Manipulation Treatment Group One-Way Anova

Affective State	Manipulation Treatment Group	Mean	Significant Difference (p value)
Positive Affect	Negative	24.64	.004*
	Positive	28.99	
	Control	26.64	
Negative Affect	Negative	16.76	.283
	Positive	15.45	
	Control	15.45	

* A post hoc multiple comparison test using Tukey HSD indicated a significant difference in positive affect between the negative and positive manipulation treatment groups.

with the findings of Westbrook (1980). Life satisfaction only displayed a significant coefficient for the multiple regression analysis using the Andrew's and Withey (1976) delighted/terrible measure of shoe satisfaction as the dependent variable. Westbrook (1980) found a significant coefficient for life satisfaction when automobile satisfaction was used as the dependent variable; no

significance of this variable was found when footwear satisfaction was used as the dependent variable. The results of the regression analyses are given in Exhibit 8.

Consideration was also given to the possible interaction effect between life satisfaction and temporary affect. Not surprisingly, a significant, positive relationship was found between life

Exhibit 6
Correlation Analysis

	Life Satisfaction	Positive Affect	Negative Affect
Life Satisfaction	1.000	.181**	-.170**
Positive Affect	.181**	1.000	.072
Negative Affect	-.170**	.072	1.000
Shoe Satisfaction (adapted from Oliver 1997)	.017	.105	.032
Shoe Satisfaction (Delighted/Terrible, Andrews and Withey 1976)	.177**	.226**	-.002
Shoe Satisfaction (Very Satisfied/Very Dissatisfied)	.109	.145*	.005
Restaurant Satisfaction (adapted from Oliver 1997)	-.014	-.017	-.068
Restaurant Satisfaction (Delighted/Terrible, Andrews and Withey 1976)	.049	-.021	-.037
Restaurant Satisfaction (Very Satisfied/Very Dissatisfied)	.034	.006	-.013

* significant at alpha = .05

** significant at alpha = .01

Exhibit 7
Non-Parametric Test
Difference in Satisfaction Scores By Manipulation Treatment Group

Scale	Treatment Group	Mean Rank	Significant Difference (p value)
Shoe Satisfaction (12 item scale adapted from Oliver 1997)	Negative	105.89	.534
	Positive	117.65	
	Control	111.21	
Shoe Satisfaction (single item: very satisfied/very dissatisfied)	Negative	99.95	.080
	Positive	122.82	
	Control	112.05	
Shoe Satisfaction (single item: delighted/terrible, Andrews and Withey 1976)	Negative	98.26	.052
	Positive	119.76	
	Control	118.19	
Restaurant Satisfaction (12 item scale adapted from Oliver 1997)	Negative	104.03	.623
	Positive	110.18	
	Control	114.08	
Restaurant Satisfaction (single item: very satisfied/very dissatisfied)	Negative	103.55	.333
	Positive	118.86	
	Control	110.70	
Restaurant Satisfaction (single item: delighted/terrible, Andrews and Withey 1976)	Negative	105.57	.350
	Positive	119.46	
	Control	108.08	

Exhibit 8
Multiple Regression Analysis: Significance of Coefficients

- Dependent Variables -

Independent Variables	SHOEDT	SHOEOSAT	SHOETOT	RESDT	RESOVSAT	RESTOT
Shoe	.000	.000	.000	.052	.063	.905
Restaurant	.201	.421	.697	.000	.000	.000
Life	.033	.520	.168	.737	.662	.436
Neg. Affect	.755	.687	.905	.201	.622	.487
Pos. Affect	.355	.164	.576	.096	.342	.561
R ²	.793	.877	.289	.915	.923	.452

Shoedt:	Andrews and Withey's (1976) delighted/terrible measure of shoe satisfaction; single item
Shoeosat:	Single item overall measure of shoe satisfaction
Shoetot:	Adaptation from Oliver's (1997) consumption satisfaction scale; 12 item measure of shoe satisfaction
Resdt:	Andrews and Withey's (1976) delighted/terrible measure of restaurant satisfaction
Resovsat:	Single item overall measure of restaurant satisfaction
Restot:	Adaptation from Oliver's (1997) consumption satisfaction scale; 12 item measure of restaurant satisfaction
Shoe:	Combined shoe satisfaction measures (combination determined by dependent variable)
Restaurant:	Combined restaurant satisfaction measures (combination determined by dependent variable)
Life:	Life Satisfaction Scale
Positive Affect:	Temporary Positive Affect Measure
Negative Affect:	Temporary Negative Affect Measure

satisfaction and temporary affect. This finding is consistent with previous research (Westbrook 1980; Peterson and Wilson 1992). However, a significant interaction was not found between life satisfaction and temporary affect in determining shoe and restaurant satisfaction ratings.

DISCUSSION

The manipulation checks revealed a significant problem with dissipation of negative affect. Although the objective of this study is not to research mood manipulation, the results reveal a need for further exploration into this area, particularly manipulation of negative affect. If manipulated negative affect can indeed dissipate within a brief ten minute period, while positive affect lingers, the results from studies in which this dissipation was not tested, are questionable.

Regression analysis indicated no influence of temporary affective state on satisfaction ratings. These results are consistent with Westbrook's (1980) study, and cause one to question the practical influence of unrelated affective state on satisfaction measurement. It is important to note, however, that the regression analyses were run

with the measures combined. It does appear that some scales are more prone to influence by affective state unrelated to the product or service being evaluated. In particular, the Andrews and Withey (1976) delighted/terrible scale appears to be susceptible.

The single item very satisfied/very dissatisfied measure of shoe satisfaction was influenced by the affective state manipulation treatment groups; however, the correlation analysis revealed that life satisfaction did not influence this measure. The adapted consumption satisfaction scale (Oliver 1997) was not influenced by either life satisfaction or temporary affect. One possible explanation for the greater sensitivity of the delighted/terrible and very satisfied/very dissatisfied measures is that these are single item scales, while the modified Oliver scale includes twelve items. Single item measures may be more susceptible to the influence of affective state unrelated to the product or service being evaluated. The varying susceptibility of the different measures is consistent with the results of Westbrook and Oliver (1991) who found that satisfaction measures vary in the extent of their relationship to the postpurchase affective response. These authors, however, were studying

consumption elicited emotions rather than unrelated affect.

Limitations

There are several limitations to this research. First, the quick dissipation of negative affect limits the ability to test the true influence of this variable on satisfaction ratings. Second, there is potential testing bias in measuring consumer satisfaction of the most recent shoe purchase and restaurant experience. The respondent could have been referring to experiences that occurred months before the measure was completed. In such situations, one may expect a larger bias from life satisfaction and affect than would be present at the time of consumption. A third limitation is the placement of the shoe and restaurant measures on the survey instrument. The instrument used in this study had all three shoe satisfaction measures listed first, then the measures of restaurant satisfaction.

CONCLUSION

This study demonstrates that life satisfaction and temporary affective state unrelated to the product or service being evaluated do sometimes influence product and service satisfaction ratings. The effect, however, is inconsistent. The Andrews and Withey (1976) delighted/terrible scale appears to be most susceptible to the influence of unrelated affective state. In this research project, affective state only had influence over the measures of shoe satisfaction; no effect was found for restaurant satisfaction ratings. This result raises questions regarding the susceptibility of product versus service evaluations.

The respondent's level of involvement with the product or service is also an issue that warrants further attention. Previous research presents the possibility that life satisfaction may have more influence on ratings of high involvement products and services. The role of involvement as a moderating factor in the relationship between temporary affect and satisfaction is a valid research concern. In this particular study it could be argued that student subjects are more highly involved with shoes – as a visible, symbolic, personal item – than with restaurants – which students on a traditional campus do not use much

and when used are primarily fast food. If involvement is indeed a moderating factor between temporary affect and satisfaction, the influence of affective state on measures of shoe satisfaction but not restaurant satisfaction ratings can be explained.

The fact that only positive affect had an influence on the ratings is also significant. Previous research, primarily in the psychology literature, has indicated inconsistent effects of negative affect. Further understanding of negative affect and its influence on the satisfaction evaluation and ratings are needed.

Although life satisfaction and temporary affect do not always influence satisfaction ratings, there are significant managerial implications, regardless. Temporary affect does show evidence of influence in certain situations. Therefore, managers must be careful to eliminate any inadvertent manipulation of mood prior to obtaining customer evaluations. If affect were systematically influenced (intentionally or not), then the ratings might be inflated or deflated and wrong conclusions drawn. Managers may also consider research to understand the naturally-occurring temporary mood states of their customers at the time of satisfaction measurement, which may vary by product or service. If mood is a big issue, then perhaps it should be accounted for in satisfaction measurement. Evidence from this research suggests that the multiple item Oliver (1997) scale is particularly robust against influence from affect and life satisfaction. Knowledge of the susceptibility of a satisfaction measure to unrelated affective state and life satisfaction is the first step toward controlling for these nuisance variables so that the true satisfaction evaluation can be obtained.

REFERENCES

- Alford, Bruce L. and Daniel L. Sherrell (1996), "The Role of Affect in Consumer Satisfaction Judgments of Credence-Based Services," *Journal of Business Research*, 37, (1), 71-84.
- Andrews, Frank M. and Stephen B. Withey (1976), *Social Indicators of Well-Being*, New York: Plenum.
- Cialdini, R. B., B. Darby and J. Vincent (1973), "Transgression and Altruism: A Case for Hedonism," *Journal of Experimental Social Psychology*, 9, 502-516.
- Cialdini, R. B. and D. T. Kenrick (1976), "Altruism as Hedonism: A Social Development Perspective on the

- Relationship of Negative Mood State and Helping," *Journal of Personality and Social Psychology*, 34, 907-914.
- Diener, Ed, Robert A. Emmons, Randy J. Larsen and Sharon Griffin (1985), "The Satisfaction With Life Scale," *Journal of Personality Assessment*, 49, (1), 71-75.
- Howard, John A. and J. N. Sheth (1969), *The Theory of Buyer Behavior*, New York: John Wiley & Sons.
- Isen, Alice (1970), "Success, Failure, Attention and Reaction to Others: The Warm Glow of Success," *Journal of Personality and Social Psychology*, 15, 294-301.
- Isen, Alice, N. Horn and D. L. Rosenhan (1973), "Effects of Success and Failure on Children's Generosity," *Journal of Personality and Social Psychology*, 27, 239-247.
- Isen, Alice, Thomas E. Shalcker, Margaret Clark and Lynn Karp (1978), "Affect, Accessibility of Material in Memory, and Behavior," *Journal of Personality and Social Psychology*, 3, (January), 1-12.
- Isen, Alice (1984), "Toward Understanding the Role of Affect in Cognition," *Handbook of Social Cognition*, R. Wyer and T. Srull, (eds.), Hillsdale, NJ: Lawrence Erlbaum Associates, 179-236.
- Mano, Haim and Richard L. Oliver (1993), "Assessing the Dimensionality and Structure of the Consumption Experience: Evaluation, Feeling, and Satisfaction," *Journal of Consumer Research*, 20, (December), 451-466.
- Mischel, W., B. Coates and A. Raskoff (1968), "Effects of Success and Failure on Self-Gratification," *Journal of Personality and Social Psychology*, 10, 381-390.
- Mischel, W., E. Eggesen, and A. Zeiss (1976), "Determinants of Selective Memory About the Self," *Journal of Counseling and Clinical Psychology*, 44, 92-103.
- Moore, B. S., B. Underwood and D. L. Rosenhan (1973), "Affect and Altruism," *Developmental Psychology*, 8, 99-104.
- Oliver, Richard L. (1997), *Satisfaction: Behavioral Perspective on the Consumer*, McGraw-Hill: New York.
- Oliver, Richard L. (1981), "Measurement and Evaluation of Satisfaction Processes in Retail Settings," *Journal of Retailing*, 57, (Fall), 25-48.
- Perdue, Barbara C. and John O. Summers (1986), "Checking the Success of Manipulations in Marketing Experiments," *Journal of Marketing Research*, 23, (November), 317-326.
- Peterson, Robert A. and William R. Wilson (1992), "Measuring Customer Satisfaction: Fact and Artifact," *Journal of the Academy of Marketing Science*, 20, 1, 61-71.
- Watson, David, Le Anna Clark and Auke Tellegen (1988), "Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales," *Journal of Personality and Social Psychology*, 54, (June), 1063-1070.
- Westbrook, Robert A. (1980), "Intrapersonal Affective Influences on Consumer Satisfaction with Products," *Journal of Consumer Research*, 7, (June), 49-54.
- Westbrook, Robert A. (1987), "Product/Consumption-Based Affective Responses and Postpurchase Processes," *Journal of Marketing Research*, XXIV, (August), 258-270.
- 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.
- Weyant, J. M. (1978), "Effects of Mood States, Costs, and Benefits of Helping," *Journal of Personality and Social Psychology*, 36, 1169-1176.
- Yi, Youjae (1990), "A Critical Review of Consumer Satisfaction," *Review of Marketing*, V.A. Zeithaml, (ed.), Chicago: American Marketing Association, 68-75.

Send correspondence regarding this article to:

Beverly K. Brockman, Doctoral Student
 University of Alabama
 Culverhouse College of Commerce and Bus. Admin.
 Department of Management and Marketing
 105 Alston Hall, Box 870225
 Tuscaloosa, Alabama 35487-0225 USA

AFFECTIVE RESPONSES TOWARDS SERVICE PROVIDERS: A CATEGORIZATION THEORY PERSPECTIVE

Rama K. Jayanti, Cleveland State University

ABSTRACT

The purpose of this article is to investigate the role of categorization processes in credence-service evaluations. Results, based on two exploratory studies, indicate that categorization theory provides a useful basis for an examination of credence-service evaluations. Further, results also demonstrate that subjects use both evaluative impressions of the service provider (derived out of categorization processes) and interaction style (perceived personal behaviors) of the service provider in judging the overall merit of the service encounter. Managerial implications and future research directions are discussed.

INTRODUCTION

Service evaluations have been traditionally examined from a disconfirmation of expectations paradigm borrowed from the product literature (Oliver 1980). The disconfirmation model may play a significant role in explaining satisfaction with services that are high on search properties or services that involve more tangible components (such as fast food restaurants). However, the model may be inadequate to explain satisfaction with service encounters that are high on experience and credence qualities for several reasons. Specific attributes are the basis for consumers' expectations, perceptions of performance, and judgments of disconfirmation in a product context. Most credence services, by definition, offer few attributes with sufficient search properties to provide much pre-purchase information. Many credence services also exhibit heterogeneity of output as a result of their high labor component. Consequently, these service encounters are not easily reduced to concrete, multi-attribute evaluations.

Faced with such information obstacles, consumers may be forced to rely on more abstract or prototypical inferences on what such services must be like. Consequently, researchers who ask for consumers' specific, attribute-based expectations may be attempting to measure items that consumers just don't use frequently in

evaluating credence-services. The use of a disconfirmation framework including pre-purchase expectations implies that consumers use cognitively derived, attribute-based expectations to judge credence services that don't supply much of that type of information.

Finally, service providers typically represent the focal point of the service for the customer (Bitner 1990). The traditional disconfirmation framework makes no provision for the consumer's affective reaction to the service provider. Oliver (1993) and Westbrook (1987) persuasively argue that satisfaction judgments should incorporate consumers' affective reactions in their composition.

The purpose of this research is to advance present understanding of service evaluations for credence type services by examining consumers' affective reactions to service providers. Specifically, we extend service evaluations literature in two directions. First, we test for the possibility of affective reactions towards the service provider preceding post-consumption evaluations by proposing categorization as an antecedent to service evaluations. Second, we test for the possibility of affective reactions towards the service provider complementing perceived performance judgements. Since the traditional disconfirmation model is well established in the literature, our interest focuses on the influences of affective reactions towards the service provider and does not include testing the traditional disconfirmation model of satisfaction. Based on the arguments put forward earlier, our interest is also centered on high credence services that involve close personal interactions. The categorization model's applicability to other types of services remains a future research possibility.

Service Encounter Evaluations

The literature on service evaluations is characterized by researchers' realization of the inherent differences between services and products and the resulting attempts to account for such differences. Service encounter evaluations have been mostly examined from a disconfirmation of

expectations perspective (Bitner 1990; Bolton and Drew 1991). Recent extensions to the disconfirmation framework included affective responses as an important component of the service satisfaction model (Oliver 1993).

Oliver (1993) proposed a model of satisfaction formation in which affect is modeled as a post-consumption process. Specifically, consumer attributions about dis/satisfaction with specific product attributes resulted in positive and negative affective reactions. In two field studies using subject evaluations of automobiles and a marketing course, support was found for the tri-component view of satisfaction as a function of cognition, affect, and direct experience.

Oliver's (1993) framework represents a significant attempt to incorporate the influence of affect on satisfaction judgments. It is noteworthy that his investigation of a service (a marketing course) revealed significantly different patterns of influence on satisfaction compared to those for the automobile judgments. Attribute-based satisfaction judgments were weakly related to overall satisfaction with the marketing course, while positive affect displayed path coefficients three times as large.

It may be possible that the informational constraints faced by respondents in generating attribute-level satisfaction judgments weakened their relationship with overall satisfaction. Within the context of credence-type service encounter evaluation, more holistic, prototypical inferences may take precedence over cognitively driven attribute-level evaluations.

In summary, although limited evidence suggests that affective evaluations play an important role in satisfaction formation, an issue to be resolved is the source of such affective reactions. We propose categorization processes to be the source of affective reactions in service situations where concrete attributes are difficult to evaluate. To explicate this possibility, we now turn our attention to a review of the categorization literature.

Categorization Processes in a Service Encounter

A growing body of literature points to the usefulness of the categorization approach in

explaining various aspects of consumer behavior (e.g., Sujan 1985; Stayman, Alden, and Smith 1992). Categorization is a simplification strategy, followed by people in an attempt to reduce complexity in their environment.

Categorization of an individual is facilitated by matching the perceived attributes of an individual to a previously stored category in consumer memory. The outcome of this process is the spontaneous transfer of affect associated with the category to the target individual. Failure to match the target individual with an accessible category may result in more attribute-oriented or "piecemeal" processing (Fiske 1982).

Service encounters, especially those services high in credence qualities, are characterized by uncertainty, ambiguity and lack of pre-purchase information (Murray 1991). More often than not, a category label is the only information available to consumers under these conditions. The ambiguity and scarcity of attribute information content in many service encounters suggests that in most instances the service provider is the service from the consumer point of view and that customer reactions towards service providers may be the most salient determinants of service encounter evaluations (Bitner 1990).

In summary, due to the informational constraints present in the service environment, we propose that categorization processes constitute the source of affective reactions towards service providers, when the target service provider fits an accessible category label. A mismatch to the available category may switch the consumer to a more attribute-based "piecemeal" processing (Fiske 1982).

Once categorization processes are established in the context of a service encounter, a possibility that merits research attention involves the role of affective reactions towards the service provider in complementing the perceived performance of the service provider. In other words, do consumers make allowances in the functional quality of the service based on their positive affective reactions to the service provider? An answer to this question may provide valuable insights to managers, since both academicians as well as managers to date have placed great importance on performance and have built their strategies around that construct alone.

In summary, the following questions specify the research expectations of the study:

- (1) Is categorization theory a helpful framework for understanding affective reactions towards service providers?
- (2) Do positive affective reactions towards service provider enhance overall service evaluations?

The first question is the focus of the first study whereas the second question was examined in study two.

STUDY 1: CATEGORIZATION PROCESSES IN SERVICE ENCOUNTERS

The present study is based on the premise that categorization is the source of affect generated towards the service provider. Limited evidence in social psychology suggests that occupation is an important category in consumers' minds (Fiske, Neuberg, Beattie, and Milberg 1987). Specifically, if there is a match between available information and category label, the affect associated with the category is spontaneously transferred to the service provider. This category-based affective reaction is termed as an "evaluative impression" in this study.

Categorization of a specific service provider is facilitated by consumers' past experiences with service providers in a specific service category (Sujan 1985). Thus, customers can quickly and easily develop an affect-laden evaluative impression of a service provider without having to judge the service encounter performance in an attribute-by-attribute basis. Since the perception of a match between a specific service provider and an accessible category in memory can take place before the service is actually purchased and/or consumed, customers can use their evaluative impression of the service provider to judge the performance of the service. In essence, the category match provides some additional information with search-like characteristics.

The research questions were examined within the context of health care services. This choice was prompted by theoretical considerations. As argued earlier, respondents' evaluative impression

of the service provider assumes importance as a determinant of service encounter evaluations when the service involved is high in experience and/or credence qualities. As the available pre-purchase cues and information content involved with a service encounter decrease, consumer reliance on heuristics should increase (Stayman et al., 1992). Accordingly, for services which require close interaction between the service provider and consumer, affect should become an important contributor to service evaluations. So, two requirements for a setting in which to test the categorization processes is that the chosen service category be high in experience and credence qualities and exhibit a strong potential for interaction between the service provider and consumer. Health care services were deemed to be appropriate to study in this context since they are high in experience and credence qualities and the typical interaction between the doctor and patient is extensive.

Procedure

The test of categorization involves extensive pretesting to establish consensual categories of interest and to assess the typical features and affect associated with the category. A typical experiment to establish categorization process involves two stages (Fiske 1982, Sujan 1985). In the first stage of the experiment, pretests are conducted to develop stimulus material and in the second stage the same stimulus material is presented to elicit the categories hypothesized in the first stage.

A series of three pretests were conducted to assess the typical features and affect associated with the category of physicians. The first issue was to establish that occupation is a potentially important category in peoples' minds and that various occupational categories elicit different affect. The second pretest was carried out to specifically test the direction of affect in the physician category.

Pretest One. Ninety undergraduate students participated in the first pretest. Two categories, physicians and lawyers were chosen to test the hypothesis that different occupational categories may elicit different affect. Half the subjects were presented with a description of a doctor and the

other half with a description of a lawyer. For purposes of this study, category-based affect was defined as a global emotional response associated with the most accessible category triggered in consumer memory. These emotional feelings are suggested to decay over time to form a generalized affective response towards the category. Past literature on categorization has operationalized category-based affect on a unidimensional scale by instructing respondents to form an impression of the target individual or a global evaluation of the individual on a single item likability scale (Fiske 1982). Keeping in view the complexity of emotions towards products proposed in consumer behavior literature (Oliver 1993; Allen, Machleit, and Marine 1988; Westbrook 1987), we used the DES (differential emotions scale) developed by Izard (1977) to measure affective responses associated with the category of physicians. Briefly, DES proposes ten primary emotions of interest, joy, surprise, sadness, anger, disgust, contempt, fear, shame, and guilt. DES has enjoyed considerable popularity in consumer research and has been reported to be a valid scale for capturing emotional responses in a product choice context (Oliver 1993; Westbrook 1987).

Subjects were instructed to think back to their past experiences with physicians (lawyers) and indicate how often they felt each of Izard's ten emotions either before, during, or after their encounter with physicians (lawyers). A set of three phrases was used to capture each of the emotion described in DES (Allen et al., 1992). For example, the emotion fear is captured by the three phrases of "feel scared, uneasy, like something might harm you"; "feel fearful, like you're in danger, very tense"; "feel afraid, shaky, and jittery". Subjects responded to each of the 30 phrases on seven point scales anchored by "never" and "very often".

The responses to the 30 phrases were summed to form an index of likability. Negative emotions were reverse scored. After confirming the unidimensionality of the scale, reliability was assessed. Both the scales had acceptable reliabilities (Cronbach's alpha for lawyers = .86; for physicians = .89). Results indicated that the physicians' category elicits significantly more positive affect compared to the lawyers' category (physicians: mean = 4.9; lawyers: mean = 2.6).

The difference between the two categories was significant ($F = 61.7, p < .01$).

Subjects were also asked, in a free elicitation task, to list attributes characteristic of and common to the category of physicians and lawyers (Sujan 1985). The salient attributes mentioned in descending order of frequency were, knowledgeable, caring, good listening skills, friendliness and sympathy. The salient attributes mentioned in descending order of frequency for the lawyer category were greedy, shrewd, aggressive, charge too much, and shifty.

The results of the first pretest provided tentative evidence that the affect associated with the lawyer category is negative whereas physicians enjoyed positive category affect. Additionally, the free elicitation task indicated that occupation is a potentially important category in subjects' minds by eliciting consensual attributes thought to be typical of the category of physicians (lawyers). To confirm these insights, a second pretest was conducted on a different sample.

Pretest Two. Sixty undergraduate students were recruited for the second pretest. Half the subjects were presented with a list of five attributes (knowledgeable, caring, good listening skills, friendliness, and sympathetic) congruent with the physician schema and the other half were presented with attributes congruent with the lawyer schema (greedy, shrewd, aggressive, charge too much, and shifty) drawn from the previous pretest. The subjects were then asked to choose among four professionals (accountant, lawyer, physician and an architect) who would ideally fit those attributes. Additionally, they were asked to read a brief description of a typical physician (lawyer) again drawn from the first pretest, and respond to a global likability scale comprising of four items (good-bad, pleasant-unpleasant, nice-awful and likable-dislikable).

The results were in general agreement with those obtained in the first pretest. The average likability of physician category was positive (mean = 2.1, below the midpoint of 3, where 1 = positive, 5 = negative), while lawyers category elicited negative affect (mean 4.2). In the physician's sample, results revealed that 79% of the subjects chose the physician, 14% chose an architect, 5% chose an accountant and 2% chose

a lawyer as an ideal description of the attributes presented. The lawyer sample overwhelmingly rated the attributes typical of a lawyer (89% lawyer, 7% accountant, and 4% architect). The results of these two pretests indicated that the subjects held consensually understood physician (lawyer) schemas and that the affect associated with the category was positive (negative).

Method

Based on the insights gained from the pretests, two videotapes were designed, one depicting the consensual attributes typical of a physician's category, and the second depicting attributes which were shown to be a mismatch to the physician's category (attributes more typical of a lawyer category). Both videotapes portrayed a spokesperson introducing himself as the marketing director of an out of town hospital group. The spokesperson then provided a verbal description of a target physician (which was a match/mismatch to the consensual attributes found in the pretest), who was under consideration to join the hospital group. The verbal description in the match condition described the physician as "working at the hospital for over eight years. He is knowledgeable, caring, and takes time to listen to his patient's problems. His patients describe him as warm, friendly, open-minded, and sympathetic. He likes to spend enough time with his patients so as to give each patient individual attention. He is highly regarded by his colleagues and enjoys a good reputation among his patients."

The verbal description in the mismatch condition described the physician as "working at the hospital for over a year. He is greedy, shrewd, and aggressive. His patients describe him as shifty, unpleasant, arrogant, and loud. He does not spend enough time with patients and is always rushed for time. He has been known to be ambitious and impersonal. His colleagues avoid him and he does not enjoy a good reputation among his patients".

The spokespersons' verbal description was accompanied by showing the subjects a photograph of the physician. The same picture was used in both the match and mismatch conditions, only the description of the physician was varied. The spokesperson then requested the subjects to

evaluate the picture along with the verbal description provided of the physician and indicate their feelings towards him on the evaluative impression scale.

The time required by subjects to respond to the evaluative impression scale (to be described shortly) was used as a measure of categorization. It was expected that subjects in the match condition would take significantly less time to provide their impression of the physician compared to their counterparts in the mismatch condition (Sujan 1985).

The videotape was pretested with another group of students who were asked how realistic, practical, and reasonable the scenario was (mean = 5.7 on a scale 1=not realistic at all to 7=very realistic). Additionally, open ended evaluations indicated that the students had no problems in relating to the scenario.

The dependent measure, evaluative impression of the physician was measured using a scale developed for this study, based on insights gained from the pretests. As indicated earlier, DES formed the basis for the measure. Since the pretests indicated a strong positive prior category affect for physicians, we included the two positive factors of interest and joy from the DES scale. Following Oliver (1993), we did not include the surprise factor because of its bivalent nature. Subjects were asked to indicate how often they experienced the two different types of emotions towards physicians on a seven point scale anchored by never and very often. The six item scale was found to be reliable with a coefficient alpha of .94. Additionally, factor analysis of the items indicated a dominant, single factor solution suggesting the unidimensional nature of the affect towards physicians in this study.

Data Collection

Data collection was facilitated by the use of a computerized questionnaire which allowed the collection of response time data. A total of 133 students (66 in the match condition and 67 in the mismatch condition), from a large southeastern university voluntarily participated in two separate computer lab sessions where they watched the videotape and completed the questionnaire. The videotape depicted the scenario described above.

Subjects were randomly assigned to groups and were instructed to indicate their feelings towards the physician on the evaluative impression scale based on the description provided. The dependent measure of interest was the evaluative impression of the physician.

Results

The effectiveness of the category match/mismatch manipulation was checked using a response time measure following Sujjan (1985). The computerized questionnaire administration facilitated the automatic recording of response times for the evaluative impression measure for each subject. There was a significant difference in processing time between the match and mismatch conditions ($F_{(1,132df)} = 27.55, p \leq .001$). Subjects in the match condition took significantly less time to provide their evaluative impression judgments compared to subjects in the mismatch condition (mean response time (match) = 1.48 min; mean response time (mismatch) = 2.09 min).

Discussion

The results of the first study found support for categorization processes in the context of service encounters. As argued earlier, for those service categories which lack concrete attributes and where attribute evaluations pose problems for consumers, the possibility exists that affective responses towards service providers precede performance evaluations. If consumers evaluate credence service providers based on their global category affect and not based on multi-attribute evaluations, is there a possibility for this category based affect to dominate overall evaluations of the same service provider? Our next study addresses this issue.

An experiment was designed to examine the complementary role of evaluative impressions in determining service provider evaluations. Evaluative impressions and perceived behaviors of the service personnel were manipulated.

STUDY 2: THE INFLUENCE OF EVALUATIVE IMPRESSIONS ON SERVICE ENCOUNTER EVALUATIONS

Research Objectives

The goal of the second study was to demonstrate the influence of evaluative impression on service encounter evaluations by manipulating evaluative impression. The ability of a positive evaluative impression of the physician to overcome mediocre "functional" (how the service was delivered) performance was of central concern in the study. The "technical" or objective performance was held constant, since there is no ecological validity to the failure of objective performance and core service failure is neither expected nor desired by the subjects. The purpose of the study was to see whether positive evaluative impression would enhance overall service evaluations even in the presence of mediocre personal behavior (termed as interaction style in this study) of a service provider. Since positive evaluative impression was elicited as a function of a match to a good physician category, we hypothesized that a match to a bad or deviant physician category may elicit negative evaluative impression. Once again, pretests were utilized to gain insights into negatively valenced category of a "bad physician". A pretest was conducted to assess subjects' perceptions regarding attributes thought to be typical of a bad physicians' category. Thirty undergraduate students were recruited for the purpose of the pretest. The subjects were requested to write down the attributes which according to them were typical of bad physicians in general, in a free elicitation format.

An analysis of the free elicitation format indicated that subjects perceive arrogance to be the most typical attribute of a bad physician. Close-mindedness, talking down to the patients, not listening to the patients problems, and over-prescribing were other typical attributes mentioned in descending order of frequency. Some of the other attributes mentioned by only one or two subjects (like unhealthiness and smoking) were eliminated. The set of attributes obtained in the present pretest (arrogant, close-mindedness, talking down to the patients, not listening to patients problems, and over-prescribing) were used

in the development of the description of a bad physician in order to elicit negative evaluative impression. The attributes used in the first study (knowledgeable, caring, good listening skills, friendliness, and sympathy) were used to elicit positive evaluative impression.

The manipulation of interaction style was based on the personal qualities of the physician. The patient satisfaction literature suggests that the perceived performance of a physician can be seen as a function of two dimensions: personal qualities and professional qualities of the physician (Hulka and Zyzanski 1982; Smith, Bloom and Davis 1985; Tucker and Tucker 1985). We term the personal qualities of the physician as "Interaction Style" of the physician which corresponds to service attributes such as friendliness, caring, and sympathy. Professional qualities of the physician correspond to service attributes such as expertise, competence, and knowledgeability. The interaction style or personal qualities of the physician was manipulated in this study.

The positive interaction style manipulation showed the physician as friendly, empathetic, and taking time to listen to patient's problems whereas the negative interaction style manipulation showed the physician as unfriendly, pushed for time, and not listening to patient's problems.

Design and Procedure

The experiment consisted of a 2 (positive vs negative evaluative impressions) X 2 (positive vs mediocre interaction style) factorial design. The stimulus development procedure followed the first study closely. The design necessitated the addition of a negative evaluative impression and interaction style manipulations to the study. Accordingly, four different videotapes were developed depicting the same scenario used in the first study.

Subjects were shown a videotape containing the sequence of events described in the first study. Briefly, the videotape portrayed a spokesperson introducing himself as the marketing director of an out of town hospital group. The spokesperson then provided a description of a target physician which was manipulated to evoke either a positive evaluative impression or a negative evaluative impression and informed the audience that he was under consideration to join the hospital group.

Specifically, the verbal description in the positive evaluative impression condition described the physician as "working at the hospital for over eight years. He is knowledgeable, caring, and takes time to listen to his patient's problems. His patients describe him as warm, friendly, open-minded, and sympathetic. He likes to spend enough time with his patients so as to give each patient individual attention. He is highly regarded by his colleagues and enjoys a good reputation among his patients."

The verbal description in the negative evaluative impression condition described the physician as "an internist at the hospital. He likes to be in-charge of the situation all the time and strongly believes that he is the only one who can make decisions about what is wrong with the patients. In the process, he usually talks "down" to his patients. He likes to overprescribe, mostly expensive medicines. He believes that most patients exaggerate their problems just to get attention. He likes to keep his patients waiting, and strongly believes that once a patient visits him, he is his property".

The spokesperson's verbal description was accompanied by showing the subjects a photograph of the physician. The same picture was used in both the positive and negative evaluative impression conditions, only the description of the physician was varied. The spokesperson then requested the subjects to evaluate the picture along with the verbal description provided of the physician and indicate their feelings towards him on the evaluative impression scale.

At this point subjects were asked to imagine a situation in which they were ill and had made an office visit to a physician with a cold, cough and flu. The videotape showed a staged interaction between the physician and a patient, portraying an office visit in which the physician gives the patient his diagnosis of the patient's condition. During the office visit, the physician appeared to be friendly, interested and concerned about the patient (positive interaction style) or pushed for time and patronizing towards the patient (mediocre interaction style). Subjects were debriefed and dismissed after collecting measures of overall service evaluation.

The videotapes were once again pretested for pragmatism, with subjects indicating that the

Table 1
Cell Means and ANOVA Tables for Service Evaluations

POS EI		NEG EI		ANOVA "F" VALUES		
POS IS	MED IS	POS IS	MED IS	EI	IS	EI x IS
4.31	1.35	3.73	1.57	.16	421.14**	4.09*

EI evaluative impression

IS interaction style

** significant p .000

* significant p .04

scenarios depicted in the videotapes reflected reasonable levels of reality (mean 5.57 on a scale from 1=not realistic at all to 7=very realistic). Subjects were undergraduate students at a large Southern University and were randomly assigned to one of the four conditions. 129 subjects (32 subjects in three cells and 33 subjects in one cell) participated in four computer lab sessions to watch a videotape of a physician patient interaction and provide their responses to a computerized questionnaire.

Measures

The experimental study involved manipulation of two independent variables (evaluative impression, and interaction style) and the dependent measure of interest was overall service evaluations. Service evaluations were measured using a five point completely dissatisfied to completely satisfied scale (are you completely dissatisfied to completely satisfied with the physician's knowledgeability, listening skill etc.), with a reliability of .83.

Evaluative impression scale was identical to the one used in the first study. An interaction style scale (alpha .96) was developed keeping in view the personal qualities explored in the past literature (listened to my problems, friendly etc.) to serve as a check for the interaction style manipulation (Smith, Bloom and Davis 1985; Tucker and Tucker 1985).

Manipulation Checks

All the manipulation check means were in the expected direction and significant differences were found across conditions. The mean score of evaluative impression manipulation check was significant ($F(1,127_{df}) = 27.05, p \leq .01$). Subjects in the positive evaluative impression group had significantly more positive perceptions of the physician compared to the negative group (positive = 3.81, negative = 3.15). The Interaction style manipulation was also successful ($F(1,128_{df}) = 456.00, p \leq .01$). Subjects in the positive interaction style condition rated the physician significantly higher on the interaction style scale compared to subjects in the mediocre condition (positive = 3.91, mediocre = 1.63).

Results

Table 1 summarizes the cell means and the analysis of variance results. The results of the experimental manipulations on subjects' perceptions of overall service with the office visit show partial support for our earlier arguments.

The results support a main effect for the interaction style manipulation on service evaluations ($F(2,127_{df}) = 421.14, p \leq .01$), suggesting that overall service evaluations differed as a function of positive vs mediocre interaction style. A main effect for evaluative impression on physician evaluations was not supported ($F(1,128_{df}) = .16, p \leq .69$). However, an interaction between evaluative impression and interaction style

was significant ($F(1,128_{df}) = 4.09, p \leq .04$) suggesting that subjects judged the merits of the service encounter based on both evaluative impression and interaction style.

Univariate tests of significance were conducted to test mean differences between treatment conditions. The mean difference between positive evaluative impression/ positive interaction style and negative evaluative impression/positive interaction style was significant ($F(1,128_{df}) = 12.15, p \leq .01$). However, the difference between positive evaluative impression/mediocre interaction style and negative evaluative impression/ mediocre interaction style conditions approached only marginal significance ($F(1,128_{df}) = 2.79, p \leq .10$). The main expectation was that positive evaluative impression would overcome mediocre interaction style of the physician. To test this hypothesis, we contrasted the positive evaluative impression/mediocre interaction style cell with the negative evaluative impression/ positive interaction style condition. Although there was a significant difference in the means they are directionally opposite to our expectations. In other words, service evaluations were higher in negative evaluative impression/positive interaction style condition compared to positive evaluative impression/mediocre interaction style condition (3.73 vs 1.35). Subjects' evaluations of the overall service were more influenced by the mediocre interaction style rather than the positive evaluative impression. However, positive evaluative impression did influence subjects' judgements when the interaction style was positive.

Discussion

The results of the study provide evidence for the importance of affective responses towards service providers in the service evaluation process. The role of affect (termed evaluative impression in this study) in physician evaluation was investigated with the help of an experimental design, where the level of affect towards the physician was experimentally manipulated. The ANOVA results found a significant interaction between evaluative impression and interaction style of the physician. However, a main effect due to evaluative impression failed to achieve statistical significance though a main effect due to interaction style was

found to be highly significant.

GENERAL DISCUSSION

The results of the two studies in this article indicate that consumers' affective reactions have a significant influence on service evaluations within the context of health care services. The findings suggest that researchers may have to incorporate the variation in service categories (Iacobucci, Grayson, and Ostrom 1994) before coming up with aggregate models of service encounter evaluations. If affective processes impact post purchase service evaluations differentially based on the level and type of attribute information available, services theory needs to incorporate this distinction. In effect, availability of attribute information may very well be the basis for a contingency model of service evaluations.

The significance of the interaction between evaluative impression and interaction style suggests that positive and negative evaluative impression exert differential effects on overall service evaluations depending on the direction of the interaction style of the physician. Under conditions of positive evaluative impression/ positive interaction style of the physician, service evaluations were elevated. However, positive evaluative impression/mediocre interaction style manipulation produced the lowest evaluation.

Positive evaluative impression/mediocre interaction style condition produced lower service evaluations compared to negative evaluative impression/mediocre interaction style condition contrary to expectations. One explanation for the counter-intuitive results may be that consumers do not like their affect expectations to be negated. In the positive evaluative impression/mediocre interaction style condition, subjects were given a description of a physician which matched their "good physician" category, following which the physician proceeded to behave in a manner which was counter to the anticipations derived out of the subject's affect. Subjects may have been more frustrated in the above situation than in a situation where they anticipated the physician to be bad based on their affect and the physician behaved in a manner which was consistent with their anticipations (negative evaluative impression/ mediocre interaction style condition). A

comparison of means in the two conditions supports such an explanation. Negative evaluative impression/mediocre interaction style condition produced higher service evaluations than positive evaluative impression/mediocre interaction style condition.

This finding points to the possibility that more damage is done by promising subjective, intangible benefits (like friendly service and empathy) and not delivering them compared to promising objective benefits (like good parking and good equipment) and not keeping those promises. From an attributional theory perspective, it may be argued that consumers make external attributions for the failure to deliver objective benefits whereas the attribution for failure to deliver subjective benefits is always internal. Consequently, consumers may be more dissatisfied with bad service than with bad parking facilities.

In summary, the positive interaction style condition produced results consistent with expectations but in the mediocre interaction style condition, the pattern of results obtained for positive and negative evaluative impression ran contrary to expectations. Interaction style of the service provider is so central to service evaluations that any kind of manipulation of interaction style of the service provider should produce a strong reaction from the consumers. Consequently, consumers may tend to discount all other information and depend solely on the mediocre interaction style to demonstrate their dissatisfaction. However, positive interaction style facilitates information processing and consumers generate enough motivation to retrieve their schematic affect to determine their level of satisfaction. Consequently, it may be advisable, at least from a theoretical point of view, to treat interaction style as the central determinant of service evaluations and investigate the antecedents to interaction style. The pattern of results obtained in this study support such an approach, since evaluative impression could explain significant variance in service evaluations only in combination with interaction style.

The results support the suggestion that evaluative impression achieves importance in service evaluations only when the interaction style is positive. As long as the interaction style of the service provider conforms to a certain threshold

level of performance predetermined by the consumers, evaluative impression achieves significance. Once this threshold level of interaction style is lowered, the lower interaction style becomes the sole determinant of service evaluations. Evaluative impression thus may be a sufficient but not a necessary condition for the determination of service encounter evaluations.

Limitations and Future Research Directions

While acknowledging the limitations of lack of generalizability due to the student sample and the simulation method used, we feel our study extends past research by incorporating schema-level affective responses as a determinant of service evaluations. Further research is clearly needed to examine the role of affective reactions in service encounter evaluation across different service categories. Additional research is also needed to extend the domain of the evaluative impression construct within the service evaluations research, and to explore the relationship between evaluative impression and information processing strategies, between evaluative impression and memory processes, and finally between evaluative impression and alternative service choice strategies.

The research reported here opens many new avenues for investigation. We hope to have kindled the interest of future researchers to further our understanding about complex constructs such as evaluative impressions since such constructs have the potential to contribute to our understanding of how consumers evaluate services with a low informational content.

REFERENCES

- Allen, Chris T., Karen A. Machleit and Susan S. Marine (1988), "On Assessing the Emotionality of Advertising Via IZARD's Differential Emotions Scale," *Advances in Consumer Research*, 15, Richard Lutz, (ed.), Ann Arbor, MI: Association for Consumer Research, 226-231.
- Bitner, Mary Jo (1990), "Evaluating Service Encounters: The Effects of Physical Surroundings and Employee Responses," *Journal of Marketing*, 54, (April), 69-82.
- Bolton, Ruth N. and James H. Drew (1991), "A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes," *Journal of Marketing*, 55, (January), 1-9.

- Fiske, Susan T. (1982), "Schema-Triggered Affect: Applications to Social Perception," *Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition*, Margaret S. Clark and Susan T. Fiske, (eds.), Hillsdale, NJ: Lawrence Erlbaum, 55-78.
- Fiske, Susan T., Steven L. Neuberg, Ann E. Beattie and Sandra J. Milberg (1987), "Category-Based and Attribute-Based Reactions to Others: Some Informational Conditions of Stereotyping and Individuating Processes," *Journal of Experimental Social Psychology*, 23, 399-427.
- Hulka, B. and S. Zyzanski (1982), "Validation of a Patient Satisfaction Scale," *Medical Care*, 20, 649-653.
- Iacobucci, Dawn, Kent A. Grayson and Amy L. Ostrom (1994), "The Calculus of Service Quality and Customer Satisfaction: Theoretical and Empirical Differentiation and Integration," *Advances in Services Marketing and Management*, Teresa A. Swartz, David E. Bowen, and Stephen W. Brown, (eds.), 3 (Greenwich, CT: JAI Press, Inc.) 1-68.
- Izard, Carroll E. (1977), *Human Emotions*, New York: Plenum Press.
- Murray, K. B. (1991), "A Test of Services Marketing Theory: Consumer Information Acquisition Activities," *Journal of Marketing*, 55, (January), 10-25.
- 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. (1993), "Cognitive, Affective, and Attribute Bases of the Satisfaction Response," *Journal of Consumer Research*, 20, (December), 418-430.
- Smith, Ruth Belk, Paul N. Bloom and Kelley Sonon Davis (1985), "Research on Patient Satisfaction: Potential Directions," *Advances in Consumer Research*, 13, Richard Lutz, (ed.), Ann Arbor, MI: Association for Consumer Research, 321-326.
- Stayman, Douglas M., Dana L. Alden and Karen H. Smith (1992), "Some Effects of Schematic Processing on Consumer Expectations and Disconfirmation Judgements," *Journal of Consumer Research*, 19, (September), 240-255.
- Sujan, Mita (1985), "Consumer Knowledge: Effects on Evaluation Strategies Mediating Consumer Judgements," *Journal of Consumer Research*, 12, (June), 31-46.
- Tucker, F. and J. Tucker (1985), "An Evaluation of Patient Satisfaction and Level of Physician Training," *Journal of Health Care Marketing*, 5, (Summer): 31-38.
- Westbrook, Robert A. (1987), "Product/Consumption Based Affective Responses and Postpurchase Processes," *Journal of Marketing Research*, 24, (Aug), 258-70.

ACKNOWLEDGEMENT

The author wishes to thank Daniel L. Sherrell for helpful comments on an earlier version of this article.

Send correspondence regarding this paper to:

Rama K. Jayanti
 Department of Marketing
 James J. Nance College of Business Administration
 Cleveland State University
 East 22 and Euclid
 Cleveland, OH 44115 USA

THE EFFECTS OF SATISFACTION AND CONSUMPTION EMOTION ON ACTUAL PURCHASING BEHAVIOR: AN EXPLORATORY STUDY

Prashanth U. Nyer, Chapman University

ABSTRACT

Studying the role that consumption emotion plays in influencing post-consumption behavior is a key to improving our understanding of the satisfaction framework. Though satisfaction research has started examining the role of affective influences, most studies view emotion merely as an antecedent to satisfaction. This study models satisfaction as a consumption emotion and hypothesizes those consumption emotions such as happiness, anger and sadness will have significant effects on both behavioral intentions and actual behavior, over and above the effects of satisfaction. This exploratory study uses data from real customers making actual purchase decisions to verify the hypothesis. Implications for satisfaction research are discussed.

INTRODUCTION

While early research on customer satisfaction concentrated on the role of cognitive influences, more recent investigations have begun looking at what role consumption emotions play in the traditional satisfaction/ dissatisfaction framework, and how consumption emotions and satisfaction might combine to influence post purchase behaviors. The growing importance of the study of emotions in this field is reflected in Woodruff's (1993) call to make the study of emotions, as it relates to CS/D&CB, a research priority. This was further highlighted by Hunt's (1993) remarks that "CS/D&CB are emotion driven, not cognition driven" and that "emotion is the critical element in CS/D&CB."

Satisfaction and Emotion

In keeping with the exploratory nature of this study, the following literature review will be limited to a few key articles. Westbrook (1987) showed that positive and negative affect were significant predictors of satisfaction over and above the traditional expectancy disconfirmation evaluations. This antecedent role of positive and negative affect was also to be found in Oliver's

(1989, 1993) models of satisfaction. Empirical evidence for the influence of affect on satisfaction also comes from Dube-Rioux (1990) and Evrard and Aurier (1994). Westbrook and Oliver (1991) argued that the common uni-dimensional satisfaction continuum coexisted with many affective experiences that influenced the judgment of satisfaction.

However, there is growing evidence that indicates that satisfaction is itself a consumption emotion, not merely a consequence of other consumption emotions. The conceptualization of satisfaction as an emotion is not new. Day (1983, p. 113) defined satisfaction as an 'emotional response manifested in feelings' while Sirgy (1984) viewed satisfaction as an emotional state resulting from the combination of cognitive evaluations. Hunt (1977, 1991) has consistently argued that consumer satisfaction is emotion and not cognition.

Various dimensional models of emotion such as the PAD model by Mehrabian and Russell (1974), the circumplex models by Russell (1980), and Watson and Tellegen (1985) include 'satisfied' as an emotional word right next to words such as 'pleased' and 'happy'. Shaver et. al. (1987) included satisfaction as a subordinate of the joy category. Bagozzi (1992) modeled satisfaction and dissatisfaction along with anger, sadness, joy, fear etc., as emotional reactions to the consumers' appraisal processes. Hausknecht (1990) in reviewing the various scales used to measure satisfaction found that satisfaction has often been measured using scales based on emotion words such as 'pleased', 'delighted' and 'contented'. Nyer (1997 a, b) found that satisfaction and happiness were too highly correlated to enjoy discriminant validity. Similarly, Babin et. al. (1994) found that satisfaction and positive affect were best modeled as one construct. All of these point to the possibility of satisfaction being an emotion. Customer satisfaction's relevance to both academicians and practitioners is because of its ability to influence post-consumption behaviors such as repurchase, W.O.M., complaining behavior and retention/ brand loyalty (see Yi 1991 for a review of some early work in this field). If satisfaction is a consumption emotion, then given

that satisfaction has significant effects on various post-consumption behaviors, do other consumption emotions such as anger and sadness also have significant influences on the post-consumption behaviors? More specifically, do these other emotions have any effects on post-consumption behaviors over and above the influence of satisfaction?

Nyer (1997 a, b) found that emotions such as anger, sadness and shame significantly contributed to the prediction of W.O.M. intentions and usage intentions, over and above the predictive ability of satisfaction. However these studies involved laboratory experiments conducted on student volunteers and therefore their external validity can be questioned. Moreover, the post-consumption behaviors studied were W.O.M. intentions and usage intentions - not actual behaviors. Since emotions are short-lived affective reactions, will the influence of consumption emotions such as happiness, anger and sadness be limited to intentions measured concurrently with the emotions, or will their effects last long enough to influence actual behaviors that occur many days later? Readers should bear in mind that it is very likely that many factors such as the involvement and experience of the consumers will influence the impact that different emotions have on various behaviors. This study attempts to take a step towards a better understanding of the role of emotion on behavior by studying the impact of a few emotions on purchasing behavior in the context of one product category. Similar studies done across many product categories, and using many different emotional constructs will be necessary to better understand the effects of emotions on behavior.

Based on previous empirical findings (Nyer 1997 a, b), it is hypothesized that:

Consumption emotions such as happiness, anger and sadness will have significant effects on post-consumption behaviors (both behavioral intentions and actual behaviors) over and above the effects of satisfaction on these behaviors.

METHOD

Subjects were 156 adults who took advantage

of a two week long promotional offer for a free one-day trial membership at a newly opened independently managed fitness center. The promotion was offered by the fitness center in an effort to increase awareness and trial. Subjects were given the questionnaire when they checked in for their one-day trial membership, and were instructed to complete the questionnaire at the conclusion of their visit to the fitness center. Completed questionnaires also served as entry forms for a drawing for two free annual memberships. After a period of three months the list of new members was examined to determine how many of the 154 participants had actually enrolled for membership (the two subjects who won the lottery for the free annual memberships were eliminated from the study).

Measures

Of the many variables included in the questionnaire, the ones relevant to this paper are the measures of satisfaction, happiness, anger, sadness and intention to enroll for membership. Subjects were instructed to mentally review their experience at the fitness center. They were then asked to indicate the extent to which they had experienced satisfaction, happiness, anger and sadness, each of which was assessed with three 7 point unipolar measures ranging from 'not at all' to 'very much'. The number of scales on the questionnaire had to be limited to ensure high response rates. Further, given the small sample size, the number of free parameters had to be restricted to ensure reasonably accurate parameter estimation. Consequently the number of emotions (apart from satisfaction) that could be included in this exploratory study was limited to three. Happiness, anger and sadness were chosen since they were distinct and commonly experienced consumption emotions.

The measures used were 'satisfied', 'dissatisfied' and 'contented' (satisfaction); 'happy', 'joyful' and 'pleased' (happiness); 'angry', 'irritated' and 'frustrated' (anger); 'sad', 'sorrowful' and 'miserable' (sadness). The measures for the consumption emotions were based on previous studies on measures of emotions (Holbrook and Batra 1987, Plutchik 1980, Richins 1997, Russell 1980 and Shaver et. al. 1987). The

'dissatisfied' scale was reverse coded before analysis. Intention to enroll was assessed using two measures. The question 'How likely are you to sign up for membership in the next few weeks?' had a 7 point response scale ranging from 'not at all likely' to 'very likely'. The question 'Do you intend to become a member soon?' required responses on a 7 point scale ranging from 'not at all' to 'very much'. Finally, actual membership enrollment behavior was measured on a dichotomous (0, 1) scale by determining if the subject had signed up for membership in the three months following the trial. Among the 38 subjects who became members (24.7% of the sample), most did so within two weeks of the trial, and none signed up for membership after 45 days. The analyses reported in the following section are based on data collected from all 154 subjects.

Analysis

The measures used in this study are ordinal and furthermore the measure of actual behavior is dichotomous. While it is usual practice to treat data measured on 5 and 7 point scales as being continuous, Joreskog and Sorbom (1996) insist that such data (including dichotomous variables) are best studied using a polychoric correlation matrix and an asymptotic covariance weight matrix. According to them, these matrices should then be analyzed with the generally weighted least-squares (WLS) option in LISREL8 rather than with the more traditional maximum likelihood (ML) estimation of the Pearson correlation matrix/covariance matrix. However the accurate calculation of asymptotic covariance matrices require larger sample sizes than were available for this study. Consequently, all the analyses in this article were conducted using both the WLS and the ML options as outlined above and the results were found to be roughly similar. While both sets of results are presented in this article, in the interest of saving space only the correlation matrix is presented in Table 1.

Results

A confirmatory factor analysis was conducted on the measures of satisfaction, happiness, anger and sadness. Table 2 represents the correlations

between the various latent variables from the Φ matrix in the confirmatory factor analysis (CFA). The latent variables satisfaction and happiness were too highly correlated ($r = 0.99$, standard error = 0.02) to exhibit discriminant validity. To formally test the hypothesis that the correlation between satisfaction and happiness is 1.0 (i.e. to test the hypothesis that they form one factor), a second CFA was conducted. This time the correlation between the two latent variables satisfaction and happiness was fixed at 1.0. Since this restricted model is nested within the previous model, a χ^2 difference test can be performed to test if the imposition of the restriction has significantly reduced the model fit (see Bollen 1989 or Hayduk 1987 for more details on the testing of nested structural equation models). The fit statistics for the full and restricted CFA models are presented in table 3. The χ^2 difference test clearly indicates that the restricted model is not significantly inferior to the full model ($p > 0.50$). In other words, fixing the correlation between satisfaction and happiness to 1 has not adversely affected the fit of the model. Thus the hypothesis that satisfaction and happiness are perfectly correlated cannot be rejected. This finding is similar to those of Nyer (1997 a, b), providing further support for the claim that satisfaction is an emotion, and perhaps a variation of joy/ happiness. Consequently, the measures of happiness and satisfaction were combined to form a new latent variable called 'Hap-Sat'. The analyses reported in this paper were also performed using the satisfaction construct (the happiness measures were excluded) and the results were very similar to what was obtained with the 'Hap-Sat' construct. In the interest of brevity, only the latter results have been reported here.

The structural equation model presented in Figure 1 was estimated. All factor loadings exceeded 0.70 and the composite reliability for the three emotional constructs were: hap-sat 0.91, anger 0.83, and sadness 0.81. The direct effect of sadness on actual enrollment was not significant in either the WLS or the ML estimation and therefore that path was dropped. Figure 1 includes both the WLS and the ML parameter. χ^2 statistics and other fit measures for this model (the full model) are provided in Table 4 and they indicate that the model in Figure 1 is an excellent representation of

Table 1
Pearson Correlation Matrix

Int1	Int2	Join	Hap1	Hap2	Hap3	Sat1	Sat2	Sat3	Ang1	Ang2	Ang3	Sad1	Sad2	Sad3
1.00														
0.61	1.00													
0.25	0.23	1.00												
0.22	0.21	0.25	1.00											
0.24	0.23	0.16	0.61	1.00										
0.33	0.24	0.18	0.65	0.67	1.00									
0.23	0.30	0.25	0.62	0.60	0.61	1.00								
0.20	0.19	0.32	0.67	0.56	0.67	0.64	1.00							
0.30	0.24	0.27	0.63	0.67	0.71	0.64	0.65	1.00						
-0.20	-0.24	-0.25	-0.25	-0.19	-0.23	-0.26	-0.19	-0.21	1.00					
-0.26	-0.19	-0.31	-0.18	-0.13	-0.21	-0.25	-0.29	-0.20	0.62	1.00				
-0.22	-0.18	-0.20	-0.22	-0.16	-0.20	-0.23	-0.16	-0.15	0.66	0.61	1.00			
-0.13	-0.10	-0.15	-0.16	-0.08	-0.07	-0.19	-0.15	-0.05	0.11	0.14	0.17	1.00		
-0.14	-0.12	-0.07	-0.10	-0.10	-0.10	-0.19	-0.07	-0.10	0.05	0.07	0.14	0.62	1.00	
-0.04	-0.05	-0.09	-0.04	-0.06	-0.05	-0.07	-0.06	-0.06	0.07	0.20	0.15	0.58	0.56	1.00

N = 154. The polychoric correlation matrix and the asymptotic covariance matrix can be obtained from the author.

Table 2
Correlation Matrix of the Latent Variables from Confirmatory Factor Analysis

	Happiness	Satisfaction	Anger	Sadness
Happiness	1.00			
Satisfaction	0.99 (0.02)	1.00		
Anger	-0.43 (0.07)	-0.45 (0.06)	1.00	
Sadness	-0.18 (0.08)	-0.17 (0.08)	0.18 (0.07)	1.00

N = 154. Data is from the WLS analysis. Data from ML estimation is similar. Numbers within parentheses represent the standard errors.

Table 3
Fit Statistics for the Full and Restricted Confirmatory Factor Analysis Models

	WLS Estimates		ML Estimates	
	Full Model	Restricted Model	Full Model	Restricted Model
χ^2 , d.f.	52.79, 48	53.02, 49	51.67, 48	51.72, 49
$\Delta\chi^2$, Δ d.f.	0.23, 1 (p > 0.50)		0.05, 1 (p > 0.70)	

the data. Since it was hypothesized that consumption emotions such as anger and sadness would be significant predictors of behavior over and above the effects of satisfaction, it is necessary to compare the full model in Figure 1 with a more restricted model in which the paths from anger and sadness to intention and actual behavior are fixed at zero. In other words, this restricted model would imply that anger and sadness have no effects on enrollment intention or actual enrollment behavior once the effect of satisfaction (and in this case the joint happiness-satisfaction construct) has

been accounted for. Since the restricted model is nested within the full model, it is possible to conduct a χ^2 difference test to examine whether fixing certain parameter values to zero has significantly reduced the model fit. The χ^2 statistics for both the full and the restricted model are presented in table 4 along with the results of the χ^2 difference test. This test of the nested models indicates that the restricted model is significantly inferior to the full model (p < 0.01).

Figure 1
Standardized Solution of the Full Model

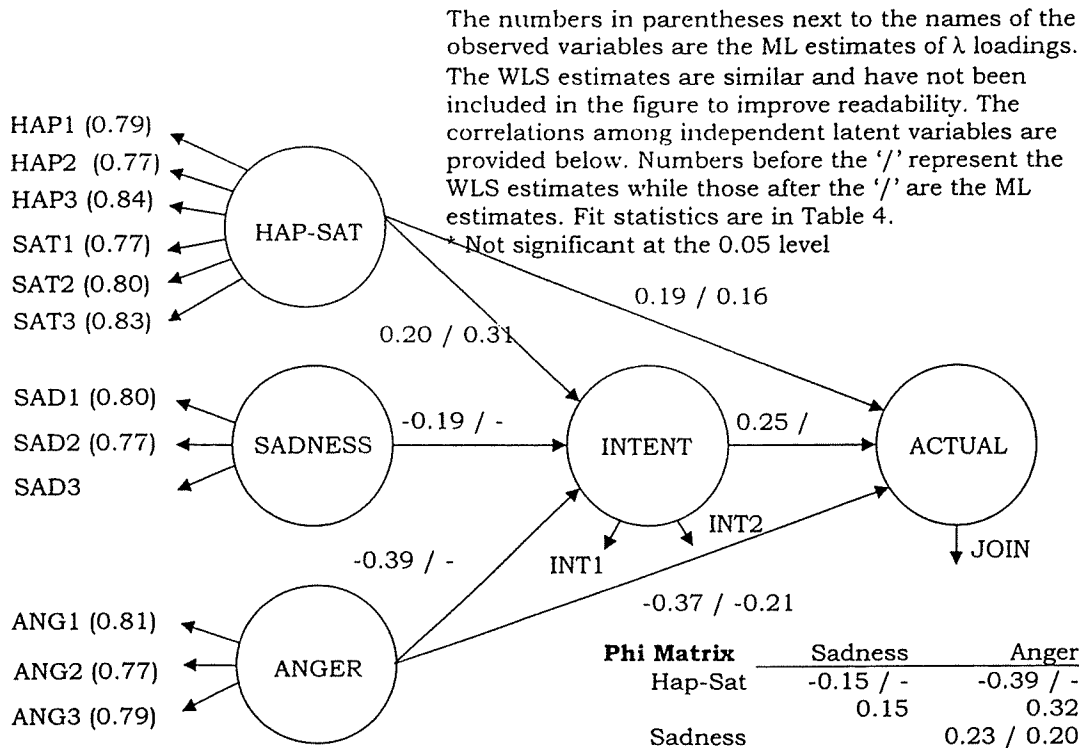


Table 4
Fit Statistics for the Full and Restricted Models

	WLS Estimates		ML Estimates	
	Full Model	Restricted Model	Full Model	Restricted Model
χ^2 , df, p	97.94, 82, 0.11	130.72, 85, 0.00	82.51, 82, 0.44	94.78, 85, 0.22
GFI, AGFI	0.98, 0.98	0.98, 0.96	0.93, 0.90	0.92, 0.89
RMSR	0.06	0.09	0.04	0.06
R ² Intent	0.33	0.30	0.21	0.16
R ² Actual	0.42	0.35	0.17	0.13
$\Delta\chi^2$, Δ d.f.	32.78, 3 (p < 0.01)		12.27, 3 (p < 0.01)	

DISCUSSION

The analysis clearly demonstrates that the full model, which includes the effects of anger, sadness and Hap-Sat (the combined happiness-satisfaction construct) on behavior intentions and actual behavior, is significantly superior to the more restricted model, which only includes the effects of Hap-Sat on the dependent variables. This

finding provides further support to the claim that post-consumption behaviors such as W. O. M., complaining behavior and repurchase should be modeled by using a broad range of emotional measures, and not just satisfaction.

Satisfaction is only one of the many different emotions that consumers experience and there is ample evidence to believe that different emotions have different effects on behaviors. For example

Lazarus (1991) provides a detailed set of action tendencies that accompany many emotions, and these action tendencies tend to be different across emotions. Thus rather than use satisfaction as an only measure of post-consumption affective reaction, researchers should use a broader range of emotional measures. However, the list of emotions that consumers experience include not only satisfaction, happiness, anger and sadness, but also fear, guilt, shame, pride, gratitude and hope among others. Does this mean that academicians and practitioners should measure all these emotions when studying post-consumption affective reactions? Clearly, not all emotions are experienced in all consumption situations. Some emotions are more common than others (Richins 1997). In the interest of parsimony, the number of emotions measured should be limited to those that are most common in a given context and to those that have the most influence on a specific behavior, such as word-of-mouth, repurchase or brand loyalty. Thus a very fruitful line of future research would be to investigate the relative influence of various emotions on the many post-consumption behaviors.

Nyer (1997 a, b) used a laboratory experiment to show that emotions had a significant effect (over and above the effect of satisfaction) on W.O.M. intentions and repurchase intentions. The present study extends Nyer's findings by showing that emotions exert their influence not only on behavioral intentions but also on actual behaviors that take place many days after the emotions are experienced. Furthermore, these findings are based on actual purchase decisions made by real consumers. However, the analysis reported in this paper is based on data from 116 'non-buyers' and a mere 38 'buyers', and as such the findings should be viewed as exploratory work. Further research using larger samples would be needed to validate these results.

REFERENCES

- Bagozzi, Richard P. (1992), "The Self-Regulation of Attitudes, Intentions, and Behavior," *Social Psychology Quarterly*, 55 (2), 178-204.
- Babin, Barry J., Mitch Griffin and William R. Darden (1994), "An Empirical Comparison of Alternative Conceptualizations of Postconsumption Reactions," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 7, 172-183.
- Bollen, Kenneth A. (1989), *Structural Equations With Latent Variables*, New York, NY: John Wiley.
- Day, Ralph L. (1983), "The Next Step: Commonly Accepted Constructs for Satisfaction Research," *International Fare in Consumer Satisfaction and Complaining Behavior*, R. L. Day and K. H. Hunt, (eds.), Bloomington, IN: Indiana University Press, 113-117.
- Dube-Rioux, Laurette (1990), "The Power of Affective Reports in Predicting satisfaction Judgments," *Advances in Consumer Research*, M. Houston, (ed.), 17, New York, Association for Consumer Research, 571-576.
- Evrard, Yves and Philippe Aurier (1994), "The Influence of Emotions on Satisfaction with Movie Consumption," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 7, 119-125.
- Hausknecht, Douglas R. (1990), "Measurement Scales in Consumer Satisfaction/Dissatisfaction," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 3, 1-11.
- Hayduk, Leslie A. (1987), *Structural Equation Modeling with LISREL*, Baltimore, MD: Johns Hopkins University Press.
- Holbrook, Morris B. and Rajeev Batra (1987), "Assessing the Role of Emotions as Mediators of Consumer Responses to Advertising," *Journal of Consumer Research*, 14, 404-420.
- Hunt, H. Keith (1977), "CS/D: Bits and Pieces," *Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, R. L. Day, (ed.), Bloomington, Division of Research, College of Business, Indiana University, 38-41.
- Hunt, H. Keith (1991), "Consumer Satisfaction, Dissatisfaction and Complaining Behavior," *Journal of Social Issues*, 47 (1), 107-117.
- Hunt, H. Keith (1993), "CS/D&CB Research Suggestions and Observations for the 1990's," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 6, 40-42.
- Joreskog, Karl and Dag Sorbom (1996), *LISREL 8: User's Reference Guide*, Chicago, IL: Scientific Software International.
- Lazarus, Richard S. (1974), *The Riddle of Man*, Englewood Cliffs, NJ: Prentice Hall.
- Lazarus, Richard S. (1991), *Emotion & Adaptation*, New York, NY: Oxford University Press.
- Mehrabian, Albert and James A. Russell (1974), *An Approach to Environmental Psychology*, Cambridge, MA: MIT Press.
- Nyer, Prashanth U. (1997 a), "A Study of the Relationships Between Cognitive Appraisals and Consumption Emotions," *Journal of the Academy of Marketing Science*, 25 (Fall) 296-304.
- Nyer, Prashanth U. (1997 b), "Modeling the Cognitive

- Antecedents of Post-Consumption Emotions," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 10, 80-90.
- 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. (1993), "Cognitive, Affective, and Attribute Bases of the Satisfaction Response," *Journal of Consumer Research*, 20, (December), 418-430.
- Plutchik, Robert (1980), *Emotions: A Psychoevolutionary Synthesis*, New York: Harper & Row.
- Richins, Marsha L. (1997), "Measuring Emotions in the Consumption Experience," *Journal of Consumer Research*, 24, (September), 127-141.
- Russell, James A. (1980), "A Circumplex Model of Affect," *Journal of Personality and Social Psychology*, 39, 1161-1178.
- Shaver, Phillip, Judith Schwartz, Donald Kirson, and Cary O'Connor (1987), "Emotion Knowledge: Further Exploration of a Prototype Approach," *Journal of Personality and Social Psychology*, 52, (6), 1061-1086.
- Sirgy, Joseph M. (1984), "A Social Cognition Model of CS/D: An Experiment," *Psychology and Marketing*, 1, (Summer), 27-44.
- Watson, David and Auke Tellegen (1985), "Toward a Consensual Structure of Mood," *Psychological Bulletin*, 98, (2), 219-235.
- Westbrook, Robert A. (1987), "Product/Consumption Based Affective Responses and Post Purchase Processes," *Journal of Marketing Research*, 24, (August), 258-270.
- 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. (1993), "Developing and Applying Consumer satisfaction Knowledge: Implications for Future Research," *Journal of Consumer satisfaction, Dissatisfaction and Complaining Behavior*, 6, 1-11.
- Yi, Youjae (1991), "A Critical Review of Consumer Satisfaction," V. Zeithaml, (ed.), *Review of Marketing*, Chicago, IL: American Marketing Association.

Send correspondence regarding this article to:

Prashanth U. Nyer
Chapman University
Orange, CA 92866 USA

THE EFFECT OF INTENSITY OF DISSATISFACTION ON COMPLAINING BEHAVIOUR

Robert Johnston, University of Warwick

ABSTRACT

Previous research has found that customers take actions, such as complain or engage in negative word of mouth behaviour, or take no action, such as never use the organisation again, based upon attitude to complaining, level of importance of the product or the nature of the problem for example. This paper seeks to extend the work on responses to dissatisfying situations by investigating the type and number of responses made against the intensity of dissatisfaction felt by the customer. Existing research has speculated about this relationship though little empirical evidence has been reported. This exploratory study is based on a convenience sample of 100 individuals who reported intensity of dissatisfaction and actions taken (rather than intentions) to dissatisfying service situations. A clear relationship is found between intensity of dissatisfaction and customer responses in terms of the types of action, the number of actions and the numbers of people told about the dissatisfying incident.

INTRODUCTION

Customers' dissatisfaction with goods and services is known to have a significant negative impact on brand loyalty and repurchase intentions (Etzel and Silverman 1981, Day 1984, TARP 1986, Singh 1990) and on the organisation's costs (Anderson et al 1994). It is important therefore that researchers and managers understand the relationship between dissatisfaction and consumer complaint behaviour (CCB). Yet the literature provides us with only limited understanding of the psychological processes that create satisfaction and dissatisfaction and yields limited information about precisely what customers might do as a result of being dissatisfied (Oliver 1997).

Several studies have attempted to identify the factors which influence the customers' responses. These antecedents of CCB include for example, the perceived likelihood of successful redress, customers' attitude to complaining, or demographics (see for example Blodgett et al

1993, Day 1984, Landon 1977, Oliver 1997). One variable that seems to have been neglected in the CCB research is the intensity of dissatisfaction felt by the customer. Whilst there has been some speculation about its relationship to CCB little empirical evidence has been reported.

This paper sets out to undertake an exploratory study to assess the relationship between customer responses and intensity of dissatisfaction. Before providing details of the study, this paper argues firstly for conducting CCB research using a wide range of responses to failure situations, secondly for treating dissatisfaction as a variable rather than a two state construct, and, thirdly it reviews the existing literature which links responses to intensity of dissatisfaction.

CONSUMER RESPONSES TO DISSATISFACTION

Prakash (1991) has criticised CCB research for developing simplistic uni-dimensional categorisations of customer responses to failure situations "While much of the conceptual research has suggested that CCB is a multi-dimensional construct, most of the empirical research has treated it as a uni-dimensional construct". (The word "responses" is used here rather than "actions" as some reactions to dissatisfaction may be in-action, such as not using the organisation again, or negative action such as negative word of mouth communication (Singh 1990).)

Examples of such uni-dimensional constructs include complain or not complain, action or no action (see for example Warland et al 1975, Morganosky and Buckley 1986). Two dimensional categorisations, however, have also been used, for example, complain or do nothing (Oliver 1997), exit or complain (Hirschman 1970, Best and Andreasen 1977). Singh (1988) went further and identified three categories of responses; voice (a verbal response directed towards a person involved in providing the service), private (negative word of mouth communication and non-use of the organisation) and third party (complain to third parties not directly involved in the service itself, such as formal or regulatory agencies or

newspapers). Singh went on to test these three categories which have been used and supported by other researchers (see for example Maute and Forrester 1993). In 1990 Singh also categorised four types of complainants; passives (people who tend to take little action, voicers (those who tend to actively complain to the service provider), irates (those people who tend to employ private responses) and activists (those with a tendency to not only complain on all three other ways but also complain to third parties).

Although most of the studies are based on a variety of responses to failure situations there seems to have been a compelling urge by many authors to compress the number of possible customer responses into a small number of broad categories. Warland et al (1975) for example identified 12 types of action from do nothing to contact a lawyer, which they reduced to two categories. Singh (1990) used ten response items based on 15 items found in the literature, from forget about the incident or do nothing through to take some legal action, which he reduced to three categories.

Whilst there appears to be agreement that customers can and do engage in varied and multiple responses (see for example Davidow and Uttal 1989, Singh 1990, Prakash 1991) this reductionist approach has several disadvantages. It loses much of the richness of the customer's responses. It limits the predictive value of the relationships found. And, importantly for this paper, it makes it difficult for managers and researchers to assess the implications of graded or escalating responses to consumer dissatisfaction.

Singh (1988) concluded that "researchers might find it advantageous to operationalise the CCB construct at the level of its individual dimensions" and that higher levels of explanation and prediction may be achieved through the operationalisation of CCB as a multi-dimensional construct.

DISSATISFACTION

Much of the work to date has also treated (dis)satisfaction as a two state construct, for example upset or not upset (Warland et al 1975), satisfied or dissatisfied (Day 1980). Prakash (1991) argued that dissatisfaction should be seen as

a variable of changing intensity. He suggested that it might be better to conceive of dissatisfaction in terms of levels or degrees of dissatisfaction. Maute and Forrester (1993) included magnitude of dissatisfaction in their study but it is unclear how the levels of dissatisfaction were operationalised except as a dichotomous construct, low dissatisfaction and high dissatisfaction. Bell and Zemke (1987) attempted to categorise levels of dissatisfaction and suggested that customers' feelings about service failures/breakdowns fall into two distinct levels of dissatisfaction which they categorised into 'annoyance' and 'victimisation'. Annoyed customers, they suggested, feel inconvenienced as the result of an experience (failure) that was slightly less than expected. A victimised customer is left with a major feeling of 'ire, frustration or pain', dependant (on the service provider) and angry. They suggested that victimised customers need to be dealt with differently to annoyed customers. Sinha (1993) describes three levels of dissatisfaction, customers who are unhappy or inconvenienced, grippers and grumblers and people who are enraged.

One of the most recent works thoroughly exploring the concept of satisfaction sheds little light on this issue. Oliver (1997) tends to use satisfaction as a two state response to the consumption of a product or service, but he also implies that it is a variable, indeed he talks about "the level of dissatisfaction" (page 20). He suggested that delight ("an expression of very high satisfaction") is an extreme expression of positive effect and disappointment a mild expression of negative effect. (It is interesting to note that he links emotions to different levels of satisfaction.) The level of (dis)satisfaction, he claimed, affects the customer's attitude to the product or service, but he does not provide more details as to what those levels might be.

Although there appears to be an emerging, though implicit, consensus that dissatisfaction is a variable there is little agreement as to what constitutes the points on the scale. Prakash (1991) suggested that researchers need to study the multi-dimensionality of dissatisfaction rather than treating it as the bipolar end of satisfaction and concluded that "This is a vast open area for research".

LINKING INTENSITY OF DISSATISFACTION TO CUSTOMER RESPONSES

This study is concerned with linking consumer responses to the intensity or levels of dissatisfaction. The idea is not new, indeed the first model proposing such a relationship was put forward by Landon in 1977. Little empirical work has followed. Prakash (1991) confirmed that the variable of intensity of dissatisfaction has not been included in studies as an independent variable and indeed that the "importance of dissatisfaction (has been) de-emphasised". Prakash (1991) went on to speculate about the relationship between the intensity of dissatisfaction and customer responses to a failure situation though he did not provide any empirical evidence.

Day (1980), on the other hand, maintained that the intensity of dissatisfaction has little value in predicting CCB outcomes implying that dissatisfaction is more of a trigger. He suggested that the nature of consumer responses is related to other factors, such as the importance of the event, customers knowledge, difficulty in seeking redress and chances of success, and demographics, lifestyle and customer values (see also Morganosky and Buckley 1986). Day (1984) maintained that evidence about the relationship between intensity of dissatisfaction and complaining behaviour is weak. It is possible that one reason for this might be that the studies in question did not attempt to use dissatisfaction as a independent variable but as one of several intervening or moderating elements in the CCB process.

More recent research has suggested that intensity of dissatisfaction may indeed have a direct effect on the customer's responses to a failure situation. Richins (1983) using severity of the problem as a surrogate for intensity of dissatisfaction found a direct relationship between intensity and complaining behaviour. Maute and Forrester (1993) concluded that the magnitude of dissatisfaction appears to be a predictor of complaining behaviour. Sinha (1993) postulated that there are links between levels of dissatisfaction and responses "... all businesses have to deal with dissatisfied customers. Such customers, however, can range from those who are simply unhappy or inconvenienced, to grippers and grumblers, to those

who are enraged, developing life-long grudges, seeking punitive action in court, or, worst of all, considering life-threatening revenge."

Prakash (1991) claimed that, while there seems to be consensus among many researchers that the intensity of dissatisfaction might be positively related to CCB, most of the work has focused on the antecedents of CCB such as attitude, expectancy value, prior experience, alienation, demographics, likelihood of successful redress, customers' attitude to complaining or the controllability of the problems, or the customers' post-complaint perception of justice (see for example Blodgett et al 1993, Day 1984, Landon 1977, Oliver 1997).

THE STUDY

This study attempts to undertake an exploratory test of the relationship between the intensity of dissatisfaction and customer responses to dissatisfying situations in order to make a small contribution to the debate on consumer responses to failure situations.

H1 The number and types of responses made by a dissatisfied customer is proportional to the intensity of the dissatisfaction

It is suggested that the more dissatisfied a customer is, the more likely s/he is to complain, to tell friends and acquaintances, to avoid using the service again and even dissuade others from using it, for example. It is suggested that mildly dissatisfied customers will initiate only a small number of responses to a failure situation unlike highly aggrieved customers who will undertake all possible responses. It is further suggested that mildly dissatisfied customers will invoke less severe responses, such as complain, whereas more dissatisfied customers will take more severe action such as actively discouraging other people from using the organisation or taking legal action for example.

H1a The intensity of dissatisfaction is directly linked to the numbers of friends and acquaintances told about the incident

One often quoted "statistic" suggests that

dissatisfied customers tell on average ten others. The TARP study (1986) for example, found that dissatisfied customers told on average nine others about their negative experience. It is suggested that the numbers told will be a function of the intensity of dissatisfaction, i.e. mildly dissatisfied customers will tell few people, highly dissatisfied customers will tell many.

Services were chosen as the focus of this study because there is a view that services can entail greater dissatisfaction than products and few studies have investigated service dissatisfaction (Singh 1990). A convenience sample of 100 individuals was chosen from the author's friends, acquaintances and colleagues, to reflect a mix of sexes and ages from 18 to 55.

Following a verbal invitation and explanation, respondents were asked to recall a recent dissatisfying experience with a service. They were provided with a simple questionnaire which asked them to "think of a time when you were dissatisfied with the level of service you received". This question allowed individuals to self select a real experience. The disadvantage of this approach might be that customers only report the more dissatisfying experiences as these may be the more memorable.

The respondents were asked to identify the type of organisation that was involved and describe in a paragraph or so what went wrong. The questionnaire asked them to identify their level of dissatisfaction at the time. Respondents were asked to report their responses following this experience. This approach has the important advantage of identifying actions taken rather than intentions which have been the focus of most studies to date, a criticism made by Singh (1988) and Prakash (1991).

Since operationalised and tested levels of dissatisfaction levels are not available in the CCB literature, these were developed in discussion with 42 executive MBA students. Initially a Likert type scale showing dissatisfaction ranging from not at all dissatisfied to extremely dissatisfied was proposed but the students found these to be unhelpful and preferred the levels described in terms of emotional outcomes. Much of the literature assumes a close relationship between levels of dissatisfaction and emotion. Oliver (1997) for example defines "delight" as "an

expression of very high satisfaction" and disappointment a mild expression of negative effect. Though he also suggested that the relationship between satisfaction and emotions is unclear, he accepts that they are closely related. Such emotional prototypes for satisfaction and dissatisfaction have been used in other studies (summarised in Oliver 1997). Oliver claimed that descriptors for "dissatisfaction (are) not so easy to pin down" but they could include anger, annoyance, frustration and hostility.

The final agreed descriptors for levels of dissatisfaction were a combination of dissatisfaction level and emotional outcomes:-

not at all dissatisfied
slightly dissatisfied
annoyed
very annoyed
extremely annoyed
absolutely furious

The responses included in the questionnaire were based upon a set of ten pre-tested items gathered from the literature by Singh 1988. These were again tested with students. The final set of responses included were:

ignore the incident (i.e. no response/do nothing)
tell friends or acquaintances
complain to the organisation
"make a fuss" with the organisation
decide never to use the organisation again
actively dissuade friends from using the service
decide actively to campaign against the organisation (explain how)
other (explain)

The "make a fuss" option was included after a small pilot study revealed, after discussion with respondents, that "to complain" was interpreted as making an official complaint to the organisation or a member of staff, i.e. "I wish to complain about ...". Several respondents explained that they did not complain as such but "made a bit of a fuss" to bring the situation to the attention of a member of staff, though not complaining per se, i.e. "I'm not complaining but ...".

Those respondents ticking the "tell friends or acquaintances" box were asked to estimate how many people they had told about the experience. The respondents were also requested to say if the organisation did anything to recover them, if so what and how they felt as a result. Respondents were invited to tick more than one box.

THE RESULTS

In previous studies much analytical work has been associated with assessing the validity of various categorisation schemes. The intention of this paper is not to validate or create such broad classifications but to assess actions taken (without further classification) against the intensity of dissatisfaction reported. Simple but powerful graphical representations of the data are provided.

Seventy seven useable questionnaires were returned. Three questionnaires which had the "not at all dissatisfied" box ticked were discarded on the basis that they were not dissatisfied customers.

Customers tended to report the more dissatisfying incidents, only six reported incidents as being slightly dissatisfying, see table 1.

Table 1
Number of Respondents and Level of Dissatisfaction Reported

Level of dissatisfaction	Number of respondents
Slightly dissatisfied	6
Annoyed	11
Very annoyed	23
Extremely annoyed	17
Absolutely furious	20
Total	77

To check if the sample was biased towards "small or large ticket" items, the responses were categorised into services estimated at over £200 in value ("big ticket" items such as car dealerships, head-hunting services and international flights) and services estimated at less than £200 ("small ticket" items such as restaurant meals, taxi services, and car maintenance). Table 2 shows that there was a fairly equal spread of low and high priced services in each of the categories.

Table 2
Estimated Value of the Service Reported

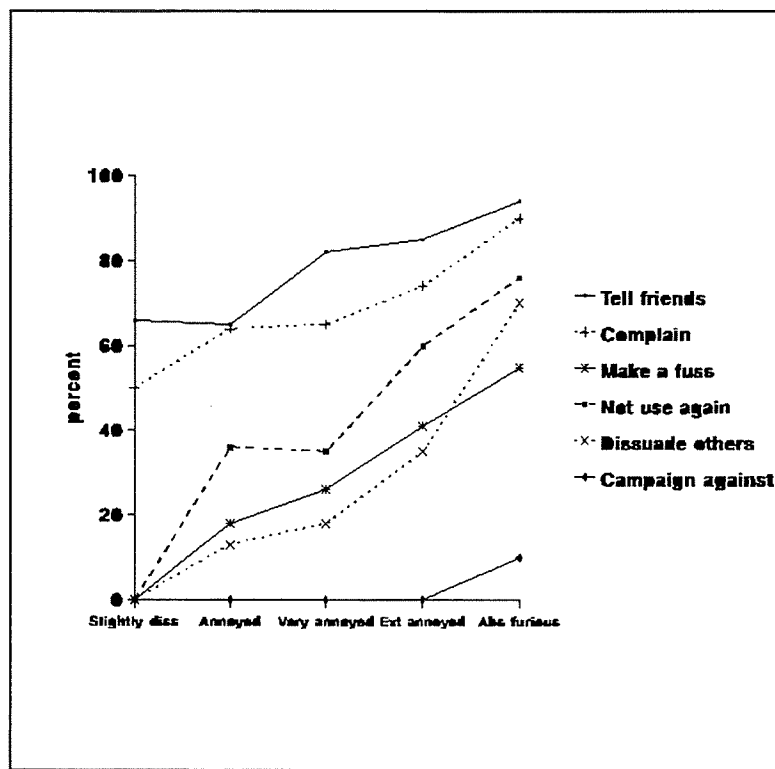
Level of dissatisfaction	Small ticket items	Big ticket items
Slightly dissatisfied	4	2
Annoyed	6	5
Very annoyed	12	11
Extremely annoyed	9	8
Absolutely furious	10	10

Figure 1 summarises in graphical form the main results of the survey by depicting the actions taken, in per cent, for each level of dissatisfaction. For example, out of the six slightly dissatisfied customers, four (66%) told friends and acquaintances about the incident and three (50%) complained to the organisation. No other actions were reported for this level of dissatisfaction. No respondent completed the "other action" line on the questionnaire, suggesting that the list of actions provided captured the key actions taken.

The findings display construct validity and behave as expected. However, a chi-squared contingency test was applied to ascertain the significance of the relationships between the actions and the level of dissatisfaction. Because of the small number of responses in some categories, some of them were merged with their adjacent categories. The null hypothesis that there is no relationship between intensity of dissatisfaction and numbers taking each action was rejected at the one per cent level for telling friends and acquaintances, actively dissuade others, and campaign against. The null hypothesis was rejected at the five per cent level for make a fuss and not use again. The null hypothesis was accepted for the complain category.

The average numbers of friends and acquaintances told about the dissatisfying incidents are shown in Figure 2. It should be noted that the higher the levels of dissatisfaction the more understated is the average. At the lower levels of dissatisfaction respondents were quite specific about the number of people they had told, however

Figure 1
Actions Taken for Different Levels of Dissatisfaction



at the higher levels several respondents reported telling 15+ or 40+ friends and acquaintances. The figures given (i.e. 15 or 40) were used to calculate the average.

DISCUSSION

Figure 1 confirms the study's main hypothesis that the number and types of responses made by a dissatisfied customer will be proportional to the intensity of the dissatisfaction with the execution of complaining which appears not to be significantly sensitive to the intensity of dissatisfaction. Every single respondent reported taking some action in response to a dissatisfying situation. The numbers of actions taken increased with each level of dissatisfaction, from an average of about two actions per slightly annoyed customer to just under four actions per absolutely furious customer. Furthermore the severity of action increased. The likelihood of not using the service again or actively

dissuading other people from using the service rises sharply with the intensity of dissatisfaction.

It is interesting to note that large numbers of only slightly dissatisfied customers were willing to take action. Although they were not willing to make a fuss at the time, and were prepared to use the service again, about half said they did tell friends and acquaintances (66%) and made a formal complaint (50%). This finding is in marked contrast to other studies (for example the TARP study 1986) which suggest that the majority of customers do not complain. On the contrary this study would suggest that the vast majority of dissatisfied customers do complain, from 50 per cent of the slightly dissatisfied customers rising to 90 per cent of the absolutely furious customers.

The majority (85%) of the absolutely furious customers told other people about the incident. Nearly all of them (90%) made a formal complaint and 55 per cent made a fuss during the service. They also took action against the organisation with

Figure 2
Average Number of People Told for Different
Levels of Dissatisfaction

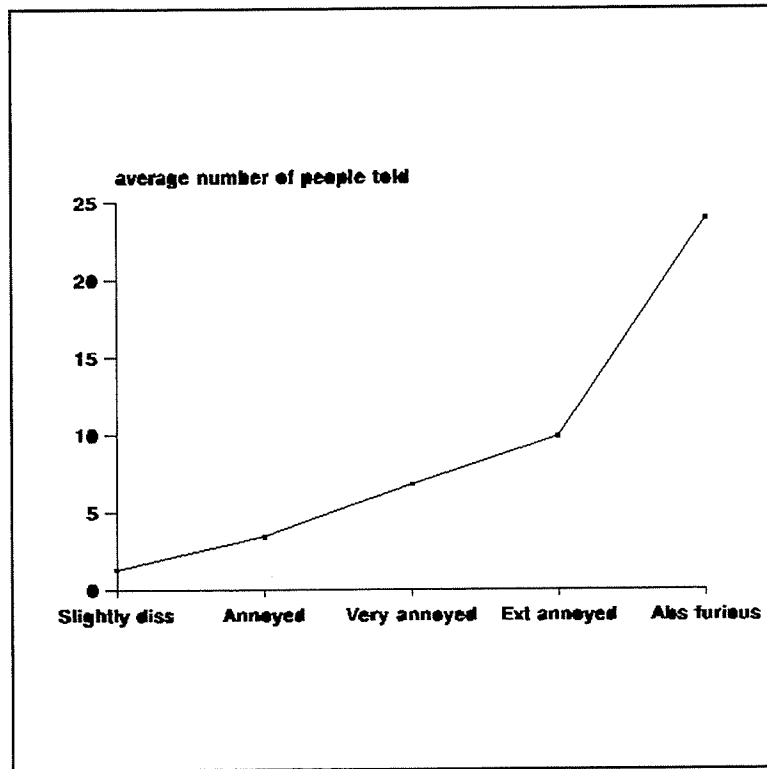


Table 3

Level of dissatisfaction	Available for recovery	Made a fuss	Complained	Both
Slightly dissatisfied	50	0	50	0
Annoyed	82	18	64	0
Very annoyed	74	26	65	17
Extremely annoyed	71	41	74	35
Absolutely furious	100	55	90	45

70 per cent of them (compared with 35 per cent of extremely annoyed customers) actively dissuading other people from using the organisation or its services. A small number, ten per cent, were prepared to go even further and actively

campaigned against the organisation, by taking out high profile legal action or petitioning outside the organisation involved.

It is important to note that the vast majority of customers made themselves available for recovery

by either making a fuss or formally complaining, or both. Table 3 summarises the percentage of respondents who made themselves available for recovery.

Fifty per cent of the slightly dissatisfied customers made themselves available for recovery by making a fuss or complaining. This rose to all the customers in the absolutely furious category making themselves known to the organisation. Clearly the majority of customers complain, in some form or other, and by doing so make themselves available for recovery. Interestingly, in the less severe cases customers tend either to complain or make a fuss, suggesting that organisations should take both equally seriously.

Figure 2 confirms the secondary hypothesis that the intensity of dissatisfaction is directly linked to the numbers of friends and acquaintances told about the incident. The numbers told rises steadily from an average of just over one in the slightly dissatisfied category to an understated average of about ten in the extremely annoyed category. This then rises sharply for the respondents who felt absolutely furious who told an understated average of over 20 people each. Interestingly the average number of people told overall is 10.1!

The number of respondents telling others also rose with the intensity of dissatisfaction. Table 4 shows the percentage of respondents in each category who reported that they told other people about the incident.

Table 4
Percentage of Respondents Who Told Other People about the Incident

Level of dissatisfaction	Percentage of respondents who told other people
Slightly dissatisfied	33
Annoyed	54
Very annoyed	69
Extremely annoyed	88
Absolutely furious	90

CONCLUSION

The study's findings need to be seen in the light of the limitations of this work. It was carried out using untested levels of dissatisfaction, though it did use a wide range of responses and collected actions rather than intentions. The sample size was very small and much greater numbers are required before any certainty could be attached to the findings.

Despite these limitations, this study has provided evidence of a strong relationship between intensity of dissatisfaction and customers' responses. It has demonstrated that the greater the intensity of dissatisfaction the more likely are customers to take "action". It has also shown that the majority of customers complain in some form or other. Eighty per cent of dissatisfied customers made themselves available for recovery by drawing the situation to the attention of staff at the time or by formally complaining.

It was found that on average 48 per cent of dissatisfied customers will not use the service again and 32 per cent will actively discourage others from using the organisation in the future. The likelihood of these responses increased with the level of dissatisfaction. Just over two per cent of dissatisfied customers will take extreme action though these are limited to the absolutely furious group. It has also been shown that on average each dissatisfied customer will tell 10 others, not a surprising finding in itself, but again the numbers told rise dramatically with the level of dissatisfaction suggesting that organisations do need to take rapid action to deal with problems before dissatisfaction escalates.

Dissatisfied customers are like a time bomb (Davidow and Uttal 1989). They can rapidly undo all the good created by large marketing budgets by spreading not only bad news but also actively mis-marketing the organisation and its services.

REFERENCES

- Anderson E. W., C. Fornell and D. R. Lehmann (1994), "Customer Satisfaction, Market Share, and Profitability: Findings from Sweden," *Journal of Marketing*, 58, July, 53-66.
- Bell C. R. and R. E. Zemke (1987), "Service Breakdown: The Road to Recovery," *Management Review*, October, 32-35.

- Best A. and A. R. Andreasen (1977), "Consumer Response to Unsatisfactory Purchases: A Survey of Perceiving Defects, Voicing Complaints, and Receiving Redress," *Law & Society Review*, 11, Spring, 701-742.
- Blodgett J. G., D. H. Granbois and R. G. Walters (1993), "The Effects of Perceived Justice on Complainants' Negative Word-of-Mouth Behavior and Repatronage Intentions," *Journal of Retailing*, 69, Winter, 399-428.
- Davidow W. H. and B. Uttal (1989), *Total Customer Service*, Harper Perennial, New York.
- Day R. (1984), "Modelling Choices Among Alternative Responses to Dissatisfaction," T. Kinnear, (ed), *Advances in Consumer Research*, 11, 496-499.
- Day R. (1980), "Research Perspectives on Consumer Complaining Behaviour," C. Lamb and P. Dunne, (eds), *Theoretical Developments in Marketing*, American Marketing Association, Chicago, 211-215.
- Etzel M. and B. Silverman (1981), "A Managerial Perspective on Directions for Retail Consumer Dissatisfaction Research," *Journal of Retailing*, 57, Fall, 124-136.
- Hirschman A. O. (1970), *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations and States*, Harvard University Press, Cambridge, Massachusetts.
- Landon E. L. (1977), "A Model of Consumer Complaint Behavior," R. Day, (ed), *Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, Bloomington Indiana, Indiana University Press, 31-35.
- Maute M. F. and W. R. Forrester (1993), "The Structure and Determinants of Consumer Complaint Intentions and Behavior," *Journal of Economic Psychology*, 14, June, 219-247.
- Morganosky M. and M. Buckley (1986), "Complaint Behaviour: Analysis by Demographics, Lifestyle and Consumer Values," Wallendorf and Anderson, (eds), *Advances in Consumer Research*, 14, 223-226.
- Oliver R. L. (1997), *Satisfaction: A Behavioral Perspective on the Consumer*, McGraw Hill, New York.
- Prakash V. (1991), "Intensity of Dissatisfaction and Consumer Complaint Behaviors," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 4, 110-122.
- Richins M. L. (1983), "Negative Word-of-Mouth by Dissatisfied Customers: A Pilot Study," *Journal of Marketing*, 47, Winter, 68-78.
- Singh J. (1990), "A Typology of Consumer Dissatisfaction Response Styles," *Journal of Retailing*, 66, 1, Spring, 57-98.
- Singh J. (1988), "Consumer Complaint Intentions and Behaviours: Definitional and Taxonomical Issues," *Journal of Marketing*, January, 93-107.
- Sinha M. (1993), "Winning Back Angry Customers," *Quality Progress*, American Society for Quality Control, November, 53-56.
- Technical Assistance Research Programs (TARP) Institute, (1986), *Consumer Complaint Handling in America: An Update Study, Part II*, performed for the US Office of Consumer Affairs, April.
- Warland R., R. Herrmann and J. Willits (1975), "Dissatisfied Customers: Who Gets Upset and Who Takes Action," *Journal of Consumer Affairs*, Winter, 148-163.

ACKNOWLEDGMENT

The author is grateful to Dr. Stephanie Stray, Warwick Business School, for her assistance with the statistical testing of the data.

Send correspondence regarding this article to:

Robert Johnston
Warwick Business School
University of Warwick
Coventry CV4 7AL
UNITED KINGDOM

EXPLAINING BRAND LOYALTY, DEALER SALES LOYALTY AND DEALER AFTER-SALES LOYALTY: THE INFLUENCE OF SATISFACTION WITH THE CAR, SATISFACTION WITH THE SALES SERVICE AND SATISFACTION WITH THE AFTER-SALES SERVICE

**Josée M. M. Bloemer, University of Limburg
Koen H. Pauwels, University of Limburg**

ABSTRACT

Building on previous research concerning brand and dealer loyalty, this study combines three satisfaction variables - satisfaction with the car, satisfaction with the sales service provided by the dealer, and satisfaction with the after-sales service of the dealer- with the three corresponding loyalty concepts. We examine whether the relationship between satisfaction and loyalty exists, which kind of satisfaction influences which kind of loyalty and how the different types of loyalty influence each other. From our study, it can be concluded that satisfaction is an important determinant of loyalty. In general one may say that the corresponding types of satisfaction (brand, sales and after-sales) predominantly determine different types of loyalty (brand, sales and after-sales). Furthermore, the different types of loyalty seem to be interdependent.

INTRODUCTION

In times of severe competition and rising customer expectations, firms are more than ever interested in keeping customers instead of attracting new ones (Heskett et al, 1994). Customer satisfaction becomes the main goal for many organizations, implicitly supposing a direct link to the final objectives like customer loyalty and profitability (Heskett et al., 1997). However, the proposed relationship between customer satisfaction and loyalty is far from obvious and still requires additional insights.

Building on previous research concerning brand and dealer loyalty (Bloemer and Lemmink, 1992), this study combines three satisfaction variables - satisfaction with the car, satisfaction with the sales service provided by the dealer, and satisfaction with the after-sales service of the dealer- with the three corresponding loyalty concepts. We examine whether the relationship between satisfaction and loyalty exists, which kind

of satisfaction influences which kind of loyalty and how the different types of loyalty influence each other. It was expected that the direct influence of each type of satisfaction on the corresponding type of loyalty would be most important (e.g. satisfaction with the car leads to brand loyalty), but that also cross-relationships could be found. Furthermore, brand loyalty, dealer sales loyalty and dealer after-sales loyalty was supposed to have a recursive relationship: dealer sales loyalty and dealer after-sales loyalty determines brand loyalty and not the other way around.

First, we present a restricted review of the literature concerning consumer satisfaction and loyalty. Secondly, we formulate our hypotheses and introduce an initial model. The research design is discussed in the next part. For both brands included in the study, we provide analyses and develop an adapted model. Finally, we conclude with managerial implications, limitations and suggestions for further research.

BACKGROUND AND REVIEW OF LITERATURE

Nowadays, the literature contains many studies about brand loyalty and repeat purchasing behavior, but the customer antecedents that determine such behavior are rarely discussed. Only by means of these antecedents, one can distinguish brand loyalty from mere repeat purchasing. A review (Bloemer et al. 1990) reveals only a few satisfaction-loyalty studies. Newman & Werbel (1973) indicate that brand loyalty appears to vary directly with the satisfaction with the old brand, but a closer look learns that the correlation is not perfect. Not all satisfied customers will be brand loyal, while not every customer who is not fully satisfied will be non-loyal. Similar results are reported by LaBarbera & Mazursky (1983), Bearden & Teel (1983), Garfein (1987), Kasper (1988), Woodside et al (1989) and Oliver & Swan (1989). A

nationwide Swedish study (Fornell, 1992; Anderson et al, 1994) even found very low correlations between customer satisfaction and loyalty for businesses like telecommunication, mail and insurances.

As indicated by Bloemer & Kasper (1995), most of the mentioned studies did not take into account the difference between repeat purchasing behavior and brand loyalty on the one hand and spurious and true brand loyalty on the other hand. When repeat purchasing behavior is studied, the focus is solely on behavior. When brand loyalty is studied the focus should be not only on the behavior but it should be as well on the attitudinal part of that behavior. Moreover, when brand loyalty is studied a distinction should be made between true and spurious brand loyalty. True brand loyalty is the repurchase of a brand based on commitment, whereas spurious brand loyalty is the repurchase of a brand not based on commitment but instead on inertia. In our research, we focus on the degree of true brand loyalty, assuming that there is a continuum between spurious and true brand loyalty based on the degree on commitment involved, and taking into account both the behavioral and the attitudinal part of loyalty.

The inter-relationship between brand loyalty and store loyalty has been the subject of a number of studies, from which it can be concluded that store loyalty is an intervening variable between satisfaction with the brand and brand loyalty (e.g. Cunningham 1956, 1961; Carman 1970; Tranberg & Hansen 1986, Bloemer et al. 1990; Bloemer and Lemmink 1992).

MODEL AND HYPOTHESES

We expand the simple model of Bloemer & Lemmink (1992) that investigated the relations between 4 constructs: satisfaction with product, satisfaction with service (sales and after-sales), (intended) brand loyalty and (intended) dealer loyalty. Our study relates satisfaction with the car (SC), satisfaction with the sales service of the dealer (SS) and satisfaction with the after-sales service (SA) to brand loyalty (BL), dealer sales loyalty (DSL) and dealer after-sales loyalty (DAL). The distinction between the sales and after-sales loyalty is important from the individual dealer's point of view. First of all, the stream of revenues

created by maintenance and repair is quite different from the profit on a one-time sale. Furthermore, a customer may have different perceptions of the sales service of the dealer (information exchange, negotiations, respectation of delivery time) and the after-sales service (repairs, payment facilities, reachability).

It might be expected that satisfaction with the car is a major positive determinant of brand loyalty. Because customers are likely to impute the satisfaction with the car to the brand. In analogy, we expect that satisfaction with the sales service is a major positive determinant of dealer sales loyalty, and that satisfaction with the after-sales service is a major positive determinant of dealer after-sales loyalty. Furthermore, interrelations between the satisfaction and the loyalty concepts might exist. For instance, satisfaction with the sales service of the dealer is likely to enhance consumer commitment towards the brand, and thus brand loyalty.

Additionally, we assume a direct positive link between dealer sales loyalty and dealer after-sales loyalty. Customers who are loyal to the dealer's after-sales service, will continue to stay in regular contact (preventive check-ups, maintenance and repair require at least bringing and picking up the car). When these customers decide to buy a new car, they are likely to consider the offer of their current dealer. Moreover, in the Dutch automobile industry it is normal practice for dealers to represent one brand only. Therefore, the customer who is loyal to the sales service of the dealer is expected to be brand loyal as well.

By the same rationale, we assume a direct positive link between dealer after-sales loyalty and brand loyalty. Since the Dutch dealer specializes in one brand, customers who want to stay loyal to the after-sales service are likely to buy the same brand of car.

Furthermore, based on the literature from the store loyalty domain, we expect that dealer sales loyalty has a direct positive link to brand loyalty. The rationale here is that especially in the Dutch automobile industry those who are loyal to their dealer are in fact, almost always, loyal to their brand also.

Moreover, based on Bloemer and Lemmink (1992), and the literature on store loyalty we expect dealer after-sales loyalty and dealer sales

loyalty to be intervening variables in the relationship between the different satisfaction concepts and brand loyalty. This means that we suppose that store loyalty leads to brand loyalty and not the other way around.

Based on the above argumentation, we formulate the following hypotheses:

H1 - Satisfaction with the car is compared with dealer sales satisfaction and dealer after-sales satisfaction the most important positive determinant of brand loyalty.

H2 - Satisfaction with the sales service is compared with car satisfaction and dealer after-sales satisfaction the most important positive determinant of dealer sales loyalty.

H3 - Satisfaction with the after-sales service is compared with car satisfactions and dealer sales satisfaction the most important positive determinant of dealer after-sales loyalty.

H4 - Dealer sales loyalty and dealer after-sales loyalty are intervening variables between satisfaction with the car, satisfaction with the sales service, satisfaction with the after-sales service and brand loyalty.

H5 - Dealer after-sales loyalty is an intervening variable between satisfaction with the sales service, satisfaction with the after-sales service and dealer sales loyalty.

Figure 1 presents the resulting model.

RESEARCH DESIGN

The respondents in the empirical part of the study are customers of different automobile dealers ($n = 407$) of two related German brands in the Netherlands. These two brands belong to the same holding, and are therefore sold by the same dealer. However, brand B is generally regarded as more exclusive and expensive than brand A. The market shares of both brands differ remarkably: brand B has a smaller share than brand A. Because the respondents had to express their feelings about the sales service, the car had to be bought less than two years ago. Furthermore, the respondents had

to have some experience with the after-sales service, which let us to impose a minimum of a one-year ownership. Previous research (Bloemer & Lemmink, 1992) found significant differences between new and used cars and between automobiles for private and for business use concerning the relationship between satisfaction and loyalty. In order to keep our study transparent, we concentrated on the new car for private use. In fact, the population for our study consists of Dutch customers, who bought a new car for private use from an official dealer of one of the two German brands between 1 and 2 years before the study. The RAI Datacentrum, where all the cars and their owners of the Netherlands are registered, provided the necessary name and address data which enabled us to draw a random sample and reach the respondents.

In 1994, 1000 owners of each brand were contacted and asked to complete a mail questionnaire. The response rate was 30.8% (615/2000). Several respondents however did not qualify to the selection criteria mentioned earlier. Moreover, because aspects of sales service were taken into account, only respondents that had 'almost always' or 'always' used the after-sales service provided by the dealer that had sold the car, were included in the analysis. Four hundred and seven out of the 615 respondents (205 and 202 for brand A and brand B respectively), constituted the final research set.

We operationalized the loyalty variables of our model as the combination of a behavioral and an attitudinal component (Jacoby & Chesnut, 1978) in accordance to the definition of Bloemer and Kasper (1995).

Brand loyalty (BL) is measured as the likelihood of repeat purchase of the brand times the degree of brand commitment. The repeat purchase measure ranged from zero (no chance at all) to ten (absolutely certain). Brand commitment is measured with a 4-item validated Likert-type 4-point scale (see Appendix).

Dealer sales loyalty (DSL) is measured as the likelihood of repeat dealer sales times the degree of dealer sales commitment. The repeat purchase measure ranged from zero (no chance at all) to ten (absolutely certain). Dealer sales commitment is measured with a 4-item validated Likert-type 4-point scale (see Appendix).

Figure 1
Model of the Relationships Between Satisfaction and Loyalty

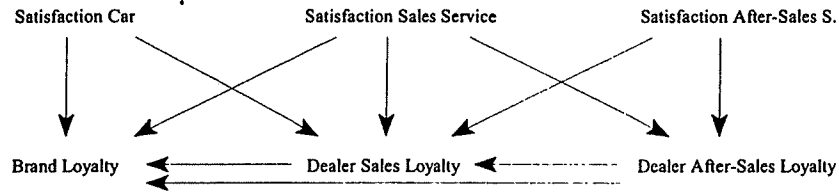


Table 1
Initial Analysis of the Variables (NS=Not Significant)
Bivariate Correlation Analysis Between Age and Loyalty Constructs

	BRAND A			BRAND B			Significance of difference
	Mean	Variance	N	Mean	Variance	N	
Satisfaction with the Car (0 - 100 %)	91.55	93.90	200	89.90	155.04	202	NS
Likelihood of Repeat Purchase Brand (0 - 10)	8.522	2.651	201	7.932	3.848	205	(P < 0.001)
Brand Commitment (4 - 16)	13.39	4.39	202	13.02	5.50	205	NS
Brand Loyalty (0 - 160)	115.44	969.67	201	105.64	1356.12	205	(P < 0.05)
Satisfaction with Sales Service (0 - 100 %)	90.60	111.07	198	89.35	145.53	199	NS
Likelihood of Repeat Dealer (3 - 10)	8.171	3.466	199	7.760	4.465	200	(P < 0.05)
Dealer Commitment Sales (4 - 16)	12.74	5.63	202	12.57	5.31	205	NS
Dealer Sales Loyalty (0 - 160)	105.85	1203.29	199	99.69	1335.21	200	NS
Satisfaction with After-Sales Service (0 - 100%)	89.75	123.67	197	88.91	153.67	200	NS
Likelihood of Repeat After-Sales (0 - 10)	9.565	0.870	200	9.525	1.541	204	NS
Dealer Commitment After-Sales (3 - 12)	9.50	3.46	202	9.33	3.59	205	NS
Dealer After-Sales Loyalty (0 - 120)	91.54	481.49	200	85.24	518.75	204	NS

Bivariate Correlation Analysis Between Age and Loyalty Constructs

AGE	Brand Loyalty	Dealer Sales Loyalty	Dealer After-Sales Loyalty	Brand Loyalty	Dealer Sales Loyalty	Dealer After-Sales Loyalty
		0.1614 (NS)	0.2640 (P < 0.001)	0.1598 (NS)	0.0489 (NS)	0.0424 (NS)

Dealer after-sales loyalty (DAL) is measured as the likelihood of repeat after-sales times the degree of dealer after-sales commitment. The repeat purchase measure ranged from zero (no chance at all) to ten (absolutely certain). Dealer after-sales commitment is measured with a 4-item validated Likert-type 4-point scale (see Appendix).

All respondents were asked what they would do in the future. In other words, only intentions are taken into account. Furthermore, the time perspective is not defined by the questions.

Satisfaction with the car (SC); satisfaction with the sales service (SS) and satisfaction with the after-sales service (SA): are measured with open-ended questions asking for the percentage of satisfaction for each.

INITIAL ANALYSIS

The initial analysis is presented in Table 1, which shows the means, variances, the significance of the difference between the scores for brand A

and brand B. It should be noticed that there is hardly any significant difference in the average scores for both brands. Exceptions are likelihood of repeat dealer sales, likelihood of repeat purchase brand and brand loyalty. Brand A shows higher scores than brand B for each of these variables.

ANALYSES FOR BRAND A

Table 2 contains an overview of the correlation coefficients between the satisfaction and the loyalty measures for brand A.

Table 2
Correlation Coefficients for Brand A Between Satisfaction Measures and Loyalty Measures

	SC	SS	SA	DSL	DAL
BL	0.51*	ns	ns	0.49*	0.26*
DSL	0.38*	0.41*	0.37*	-	0.40*
DAL	0.27*	0.30*	0.41*	0.40*	-

* Correlation coefficients significant at $P < 0.001$

High positive correlations exist for the direct links between each type of satisfaction and the corresponding type of loyalty. Also cross-relationships are demonstrated, like the positive relation between satisfaction with the after-sales service (SA) and dealer sales loyalty (DSL). Furthermore, interdependence between the different types of loyalty could be expected because of the high correlation coefficients.

In order to gain additional insight into the net contribution of the satisfaction measures with respect to the loyalty measures, we computed partial correlation coefficients (Table 3).

From Table 3 it can be seen that there is a significant positive partial correlation between the satisfaction with the car and brand loyalty. However, there is also a significant positive correlation between satisfaction with the car and both dealer sales loyalty and dealer after-sales loyalty. Nevertheless, these results are in agreement with our first hypothesis that the most important determinant of brand loyalty is the satisfaction with the car. The first hypothesis can be accepted. The same reasoning can be applied

for satisfaction with the sales service and satisfaction with the after-sales service in relation to sales loyalty and after-sales loyalty respectively. Therefore, also hypotheses 2 and 3 can be confirmed.

Table 3
Partial Correlation Coefficients Between Loyalty Measures and Satisfaction Measures (Controlled for the Other Satisfaction Measures)

	SC	SS	SA
BL	0.52	-0.00	-0.11
DSL	0.26	0.23	0.07
DAL	0.11	0.06	0.25

In order to draw more definite conclusions, the general model as presented in Figure 1 with different proposed restrictions, will be tested.

General model

$$BL = \beta_{11} + \beta_{12} DSL + \beta_{13} DAL + \beta_{14} SC + \beta_{15} SS + \beta_{16} SA + \epsilon_1$$

$$DSL = \beta_{21} + \beta_{22} BL + \beta_{23} DAL + \beta_{24} SC + \beta_{25} SS + \beta_{26} SA + \epsilon_2$$

$$DAL = \beta_{31} + \beta_{32} BL + \beta_{33} DSL + \beta_{34} SC + \beta_{35} SS + \beta_{36} SA + \epsilon_3$$

Model restriction number	Estimation method	Restrictions
1 DSL & DAL are mediating variables	Recursive system OLS	$\beta_{22}, \beta_{23} = 0$ $\beta_{33} = 0$
2 DAL does not explain BL SS & SA do not explain BL SA does not explain DSL mutual dependence between BL and DSL	Two-stage-least squares	$\beta_{13} = 0$ $\beta_{15}, \beta_{16} = 0$ $\beta_{26} = 0$
3 SA does not explain DSL SS does not explain DAL BL does not explain DAL mutual dependence between DSL and DAL	Two-stage-least squares	$\beta_{26} = 0$ $\beta_{35} = 0$ $\beta_{32} = 0$

As mentioned above, the first model we test presumes that dealer sales loyalty and dealer after-sales loyalty are seen as mediating concepts between the satisfaction measures and brand loyalty and that dealer after-sales loyalty is a mediating variable between satisfaction with the sales, after-sales service and dealer sales loyalty. This leads to the following restricted model formulation (model 1A):

Model 1A

$$\begin{aligned} BL &= f(SC, SS, SA, DSL, DAL) \\ DSL &= f(SC, SS, SA, DAL) \\ DAL &= f(SC, SS, SA) \end{aligned}$$

The results of the path analysis, on the basis of this model, are presented in Table 4.

Most of our earlier conclusions can be confirmed by this analysis. Satisfaction with the car is the most important determinant of brand loyalty, satisfaction with the sales service is the most important determinant of dealer sales loyalty and satisfaction with the after-sales service is the most important determinant of after-sales service. Furthermore, significant positive relationships are found between dealer sales loyalty and brand loyalty on the one hand and dealer after-sales loyalty and dealer sales loyalty on the other hand. However, the direction of the relationship between these concepts is not yet determined upon. The possibility of mutual dependence will be examined by means of the two-stage-least square method. The correlation between dealer after-sales loyalty and brand loyalty is not significant and will be therefore be discarded in further discussion.

Although the foregoing analysis did provide more insight into the data, the possible mutual

relationships between the loyalty variables was not yet taken into account. In order to do so we may estimate the following models:

$$\begin{aligned} BL &= f(SC, SS, SA, DSL, DAL) \\ DSL &= f(SC, SS, SA, BL, DAL) \\ DAL &= f(SC, SS, SA, BL, DSL) \end{aligned}$$

However, this set of equations is over-identified (Johnston 1972, Naert & Leeflang 1978), hence the parameters could not be estimated. Because our previous analysis showed that satisfaction with the sales service, satisfaction with the after-sales service and dealer after-sales loyalty do not have a significant impact on brand loyalty and that the influence of satisfaction with the after-sales service on dealer sales loyalty is not-significant, we formulate an alternative model, which tests the interdependence between brand loyalty and dealer sales loyalty:

Model 2A

$$\begin{aligned} BL &= f(SC, DSL) \\ DSL &= f(SC, SS, BL, DAL) \end{aligned}$$

Table 5 shows the estimated parameters for this model.

It can be concluded that satisfaction with the car has a positive impact on brand loyalty and that satisfaction with the sales service positively influences dealer sales loyalty. Moreover, brand loyalty and dealer sales loyalty is obviously mutually dependent, as indicated by the almost similar beta coefficients. Therefore, hypothesis 4 can not be confirmed: dealer sales loyalty and dealer after-sales loyalty are not intervening variables between the satisfaction concepts and

Table 4
Beta Coefficients for Model 1A Based on Regression Analysis

	SC	SS	SA	DSL	DAL	R ²
BL	0.44*	ns	ns	0.41	ns	0.42
DSL	0.23*	0.25*	ns	-	0.26*	0.30
DAL	ns	ns	0.32*	-	-	0.10

* Coefficients significant at $P < 0.001$

Table 5
Two-Stage-Least-Square Analysis of Model 2 A; Beta Coefficients

	SC	SS	BL	DSL	DAL	R ²
BL	0.38*	-	-	0.35*	-	0.38
DSL	ns	0.31*	0.40*	-	0.19*	0.43

* Coefficients significant at $P < 0.001$

Table 6
Two-Stage-Least-Square Analysis of Model 3 A; Beta Coefficients

	SC	SS	SA	DSL	DAL	R ²
DSL	0.23*	0.26*	-	-	0.26*	0.30
DAL	-	-	0.30*	0.29*	-	0.23

* Coefficients significant at $P < 0.001$

Figure 2
Final Mode Brand A



brand loyalty. However, brand loyalty and dealer sales loyalty are mutually dependent. These findings are in contrast with the conclusions of Bloemer & Lemmink (1992), who reported a one-

directional impact of dealer (sales) loyalty on brand loyalty.

As noted above, we also have to test the possible mutual dependence between dealer sales

loyalty and dealer after-sales loyalty. On the basis of previous findings, the following restricted model is estimated:

Model 3A

$$DSL = f(SC, SS, DAL)$$

$$DAL = f(SA, DSL)$$

Table 6 shows the estimated parameters for this model.

Again, the mutual dependence of two loyalty measures is demonstrated in Table 6. Hypothesis 5 can not be confirmed. Dealer after-sales loyalty is not an intervening variable between satisfaction with the sales service, satisfaction with the after-sales service and dealer sales loyalty.

Our final model for brand A can be formulated as follows:

$$BL = \beta_{11} + \beta_{12} DSL + \beta_{13} SC + \epsilon_1$$

$$DSL = \beta_{21} + \beta_{22} BL + \beta_{23} DAL + \beta_{24} SS + \epsilon_2$$

$$DAL = \beta_{31} + \beta_{32} DSL + \beta_{33} SA + \epsilon_3$$

This model is also depicted in Figure 2.

In summary, hypotheses 1, 2, and 3 can be confirmed: a direct link between each satisfaction concept and the corresponding loyalty concept appears to exist. Satisfaction with the car is the most important determinant of brand loyalty, satisfaction with the sales service is the most important positive determinant of dealer sales loyalty, and satisfaction with the after-sales service is the most important positive determinant of dealer after-sales loyalty. The mutual dependence between the loyalty constructs however, resulted in a rejection of hypotheses 4 and 5, which specified a non-recursive relationship between the different types of loyalty.

ANALYSES FOR BRAND B

The analysis for brand B, generally regarded as the more exclusive brand, shows different results.

First of all, all correlation coefficients are significant in the analysis for brand B (Table 7). In contrast to the findings for brand A, also satisfaction with the sales service and satisfaction

with the after-sales service seem to have a positive impact on brand loyalty. The direct positive influences of satisfaction with the car on brand loyalty and of satisfaction with the sales service on dealer sales loyalty, show almost similar coefficients as the ones found for brand A. However, satisfaction with the after-sales service does not yet seem to be the most important determinant of dealer after-sales loyalty. Furthermore, the correlation coefficients between the different types of loyalty for brand B are even higher than those found for brand A.

The partial correlation coefficients between the different variables for brand B are shown in Table 8.

Table 7
Correlation Coefficients for Brand B Between Satisfaction Measures and Loyalty Measures

	SC	SS	SA	DSL	DAL
BL	0.52*	0.29*	0.29*	0.71*	0.44*
DSL	0.40*	0.40*	0.33*	-	0.50*
DAL	0.29*	0.23*	0.23*	0.50*	-

* Correlation Coefficients significant at $P < 0.001$

Table 8
Partial Correlation Coefficients Between Loyalty Measures and Satisfaction Measures (Controlled for the Other Satisfaction Measures)

	SC	SS	SA
BL	0.43	0.07	-0.01
DSL	0.26	0.21	-0.00
DAL	0.20	0.06	0.03

In correspondence with Table 3, there is a significant positive partial correlation between the satisfaction with the car and brand loyalty. There also is a significant positive partial correlation between satisfaction with the car and both dealer sales loyalties as dealer after-sales loyalty. A

significant positive link between satisfaction with the sales service and the dealer sales loyalty is also demonstrated. Based on Table 8, hypotheses 1 and 2 can be confirmed for brand B, whereas hypothesis 3 has to be rejected.

Next, model 1B is tested for brand B. The results are presented in Table 9.

Model 1B

$$BL = f(SC, SS, SA, DSL, DAL)$$

$$DSL = f(SC, SS, SA, DAL)$$

$$DAL = f(SC, SS, SA)$$

The most striking difference with the coefficients for brand B in Table 9, is the non-significance of the influence of satisfaction with the after-sales service on dealer after-sales loyalty.

Again here, the dependence between the loyalty constructs is examined by means of the two-stage-least square method. Based on the previous results, we formulate the following model:

Model 2B

$$BL = f(SC, DSL)$$

$$DSL = f(SC, SS, BL, DAL)$$

Table 10 shows the results of the analysis of model 2B.

In accordance to our previous findings, hypotheses 1 and 2 can be confirmed for brand A: satisfaction with the car has a positive impact on brand loyalty and satisfaction with the sales service positively influences dealer sales loyalty. Interestingly, the influence of satisfaction with the car on dealer sales loyalty seems to disappear when the mutual dependence between brand loyalty and dealer sales loyalty is incorporated. Cross-relationships between satisfaction and loyalty concepts are felt through the strong mutual relationship between brand loyalty and dealer sales loyalty. Hypothesis 4 can therefore not be confirmed.

By analogy with the analysis of brand A, we also examine the mutual dependence of dealer sales loyalty and dealer after sales loyalty for brand B. Because of the non-significance of satisfaction with the after-sales service and dealer after-sales

loyalty, the formulation of Model 3B will differ from that of Model 3A:

Model 3B

$$DSL = f(SC, SS, DAL)$$

$$DAL = f(SC, DSL)$$

The results of the analysis of model 3B are provided in Table 11.

As is the case for brand A, hypothesis 5 can neither be confirmed for brand B (the mutual dependence of the loyalty measures is shown by their beta coefficients, which are in line with one another).

In summary, the analysis for brand B demonstrates differences as compared to with the findings for brand A. Although hypotheses 1 and 2 stand the test, hypothesis 3 can not be confirmed. Satisfaction with the after-sales service does not appear to have a significant positive impact on dealer after-sales loyalty. These findings enforce our doubt about an "obvious" relation between satisfaction and loyalty in all cases. Concerning the mutual dependence between the loyalty constructs, we may conclude for brand B as well as for brand A that hypothesis 4 and 5 has to be rejected. There is supposed to be a recursive relationship between the loyalty constructs. Our final model for brand B can be formulated as follows:

$$BL = \beta_{11} + \beta_{12} DSL + \beta_{13} SC + \epsilon_1$$

$$DSL = \beta_{21} + \beta_{22} BL + \beta_{23} DAL + \beta_{24} SS + \epsilon_2$$

$$DAL = \beta_{31} + \beta_{32} DSL + \epsilon_3$$

This model is depicted in Figure 3.

CONCLUSION AND MANAGERIAL IMPLICATIONS

We presented a model with respect to the relations between three satisfaction constructs and three loyalty constructs and tested it with data collected for two data sets: brand A and brand B. From this study, we can conclude that satisfaction is an important determinant of loyalty. In general, one may say that the corresponding types of satisfaction (brand, sales and after-sales)

Table 9
Beta Coefficients for Model 1B Based on Regression Analysis

	SC	SS	SA	DSL	DAL	R ²
BL	0.31*	ns	ns	0.58*	ns	0.58
DSL	0.19*	0.23*	ns	-	0.40*	0.36
DAL	0.23*	ns	ns	-	-	0.10

* Coefficients significant at P < 0.001

Table 10
Two-Stage-Least-Square Analysis Model 2B; Beta Coefficients

	SC	SS	BL	DSL	DAL	R ²
BL	0.28*	-	-	0.59*	-	0.56*
DSL	ns	0.21*	0.58*	-	0.19*	0.57*

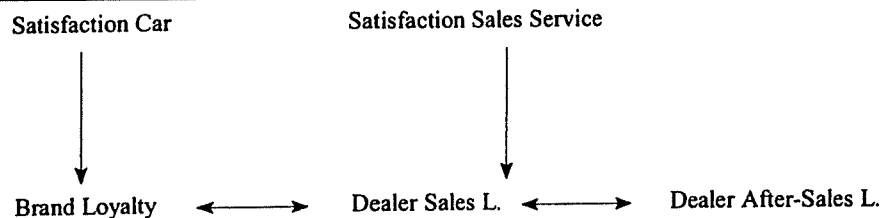
* Coefficients significant at P < 0.001

Table 11
Two-Stage-Least-Square Analysis of Model 3 B; Beta Coefficients

	SC	SS	DSL	DAL	R ²
DSL	0.21*	0.23*	-	0.37*	0.36
DAL	ns	-	0.45*	-	0.23

* Coefficients significant at P < 0.001

Figure 3
Final Model Brand B



predominantly determine different types of loyalty (brand, sales and after-sales). Furthermore, it can be concluded that the different types of loyalty are interdependent. Specifically, for brand A,

hypotheses 1, 2 and 3 were confirmed: a direct link between each satisfaction measure and each loyalty measure appears to exist. The mutual dependence between the loyalty constructs

however, resulted in a rejection of hypotheses 4 and 5, which specified a non-recursive relationship between brand-, sales- and after-sales loyalty. In the case of the more exclusive brand B, the positive relation between satisfaction with the after-sales service and dealer after-sales loyalty could not be confirmed. In addition to hypotheses 4 and 5, who were also not confirmed, also hypothesis 3 has to be rejected too. For both brand A and brand B we proposed adopted models which are in line with the findings of our research.

A possible explanation for the different findings for brand A and brand B, can be found in the exclusive image of brand B. We might argue that the customers of brand B have a stronger bond with their brand and their dealer. In other words, these customers may perceive a stronger connection between their personal values/way of living and their ownership of the car. Therefore, it does not matter that much whether they are extremely satisfied or just satisfied on average with the after-sales service.

An alternative interpretation could be the following: brand B is also a popular car for companies to lease. These users of the car are employees who do not normally have to pay for or experience the after-sale service. Therefore dealer after-sales loyalty might not be significantly influenced by the after-sales satisfaction. Further research is needed to examine these interpretations.

Various general implications of this study can be formulated for both the car manufacturer and the dealer. The analysis of the generalized data lead us to advise the manufacturer to (1) strictly monitor the quality of the physical product in order to gain high customer satisfaction, which has a positive impact on brand loyalty for both brands included in our study, (2) ensure that excellent sales service is provided by the dealer, because of its positive influence on brand loyalty via the mutual dependence with dealer sales loyalty; and (3) ensure excellent after-sales service, especially for brands like brand B, for which the respondents used their evaluation of the after-sales-service in their dealer loyalty decision (which has an indirect impact on brand loyalty via dealer sales loyalty).

From the dealer's point of view, our analysis can help to decide on the focus in service and communication to owners of brand A and brand B.

This leads to specific implications of the study. Because dealer after-sales loyalty seems to be linked rather to dealer sales loyalty than to satisfaction with the after-sales service for brand B, it could be rewarding to develop a personal relation with the customer. The dealer could give him/her special attention, inform him/her about new models that are (going to be) introduced and enforce the exclusive and high-quality perception of the brand and its dealer network. The importance of contact personnel in creating the customer feeling that he/she is special, should not be underestimated. Although a qualitative after-sales service is still important, special efforts to increase customer satisfaction with this service will not be very effective in raising dealer loyalty. In contrast, owners of brand A are more likely to evaluate the after-sales service thoroughly. In this case, the dealer would be rewarded in terms of loyalty for delighting the customer with excellent maintenance and repair service. Dealer personnel that are perceived by customer as excellent service providers, are likely to enhance dealer after-sales and dealer sales loyalty. In general the results of our study lead us to advise to the dealer to (1) provide optimal service (sales and after-sales) in order to assure dealer loyalty, (2) in line with the previous point: insist on the support of the manufacturer to help to provide perfect service (3) guarantee an excellent quality product in order to gain dealer loyalty through brand loyalty; and (4) not to underestimate the importance of the contact personnel in creating loyalty via their behavior and their way of rendering various services.

Limitations

First, the most important limitations of this study concern the brand included and the sample that took part in this study. It should be mentioned that the context of car distribution in the Netherlands, due to exclusiveness, is quite unique. Second, the sample is somewhat biased towards loyal behavior, because consumers are chosen only in so far as they have been servicing their car from the dealer they have bought it from. This might have reduced the variability in dealer (after-)sales loyalty and the generalizability of our findings. Third, all our measures are paper and pencil

measures and we lack measures of actual repeat buying behavior. Moreover the time frame is important here. Does the consumer intend to buy another car next month, next year or in another five to ten years. Fourth, the proximity of the questions might have led to cognitive consistency patterns and other demand characteristics or response biases. It could be expected that this favored the applicability of our model.

Suggestions For Further Research

A number of questions remains open from this study and needs to be addressed in further research. One further question concerns the reasons for the differences we found between the two brands. More detailed inquiry into the differences between a low priced and a premium priced brand is needed. Furthermore, in future research, samples should no longer be limited to customers who patronize the same dealer as they bought their car from. This study should be repeated in other market settings where one dealer is not restricted to selling brands from only one manufacturer. Of course, actual buying behavior of customers on a longitudinal basis should be taken into account. In addition, attention should be paid to the competitive actions of car dealers and manufacturers of other brands. Finally, the product sample needs to be expanded to low involvement products.

REFERENCES

- Anderson, E., C. Fornell and D. Lehmann (1994), "Customer Satisfaction, Market Share and Profitability: Findings from Sweden," *Journal of Marketing*, 58, 53-66.
- Bearden, W. O. and J. E. Teel (1983), "Selected Determinants of Consumer Satisfaction and Complaint Reports," *Journal of Marketing Research*, 10, 309-313.
- Bloemer, J. M. M. (1993), *Loyaliteit en tevredenheid* (Loyalty and satisfaction), Datawyse, Maastricht.
- Bloemer, J. M. M. and H. D. P. Kasper (1995), "The Complex Relationship Between Consumer Satisfaction and Brand Loyalty," *Journal of Economic Psychology*, 16, 311-329.
- Bloemer, J. M. M., J. D. P. Kasper and J. G. A. M. Lemmink (1990), "The Relationship Between Overall Dealer Satisfaction, Satisfaction with the Attributes of Dealer Service, Intended Dealer Loyalty and Intended Brand Loyalty: A Dutch Automobile Case," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 3, 42-47.
- Bloemer, J. M. M. and J. G. A. M. Lemmink (1992), "The Importance of Customer Satisfaction in Explaining Brand and Dealer Loyalty," *Journal of Marketing Management*, 8, 351-364.
- Carman, J. M. (1970), "Correlates of Brand Loyalty; Some Positive Results," *Journal of Marketing Research*, 7, 67-76.
- Cunningham, R. M. (1956), "Brand Loyalty - What, Where, How Much?" *Harvard Business Review*, 34, 116-128.
- Cunningham, R. M. (1961), "Customer Loyalty to Store and Brand," *Harvard Business Review*, 39, 127-128.
- Fornell, Claes (1992), "A National Customer Satisfaction Barometer: The Swedish Experience," *Journal of Marketing*, 56, 6-21.
- Garfein, R. (1987), "Evaluating the Impact of Customer Service Delivery Systems," *Marketing Review*, 42, 11-15.
- Heskett, J. L., T. Jones, G. Lovemer, W. Sasser and L. Schlesinger (1994), "Putting the Service-Profit Chain to Work," *Harvard Business Review*, 164-171.
- Heskett, J. L., W. E. Sasser and L. A. Schlesinger (1997), *The Service Profit Chain: How Leading Companies Link Profit and Growth to Loyalty, Satisfaction and Value*, The Free Press, New York.
- Jacoby, J. and R. W. Chesnut (1978), *Brand Loyalty Measurement and Management*, New York, John Wiley and Sons.
- Johnston, J. (1972), *Econometric Methods*, 2nd ed. McGraw-Hill, New York.
- Kasper, J. D. P. (1988), "On Problem Perception, Dissatisfaction and Brand Loyalty," *Journal of Economic Psychology*, 9, 387-397.
- LaBarbera, P. A. and D. Mazursky (1983), "A Longitudinal Assessment of Consumer Satisfaction/Dissatisfaction," *Journal of Marketing Research*, 20, 393-404.
- Naert, P. A. and P. S. H. Leeflang (1978), *Building implementable marketing models*, Martinus Nijhoff Social Sciences Division, Leiden/Boston.
- Newman, J. W. and R. A. Werbel (1973), "Multivariate Analysis of Brand Loyalty for Major Household Appliances," *Journal of Marketing Research*, 10, 404-409.
- Oliver, R. L. and J. E. Swan (1989), "Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions: A Field Survey Approach," *Journal of Marketing*, 53, 21-35.
- Tranberg, H. and F. Hansen (1986), "Patterns of Brand Loyalty: Their Determinants and Their Role for Leading Brands," *European Journal of Marketing*, 30, 81-109.
- Woodside, A. G., L. L. Frey and R. T. Daly (1989), "Linking Service Quality, Customer Satisfaction and

Behavioral Intention," *Journal of Health Care Marketing*, 5-17.

APPENDIX

Brand commitment (dealer sales commitment en dealer after-sales commitment) were measured with the following 4 items:

- When another brand of car is on sale, I will purchase it, rather than my preferred car brand.
- If my preferred car brand were not available right away, it would make little difference to me if I had to choose another brand.
- If my preferred car brand is not available, I will buy another favorite brand.
- If I have to make a choice for a particular car brand before actually making the purchase, I might easily change my intended choice upon receiving discrepant information.

By substituting the words car brand of brand of car, comparable questions can be formulated for the dealer sales service and the dealer after-sales service.

The unweighted items were combined in a score per respondent. The minimum being 4 and the maximum score being 16 on the 4-point Likert scale. Cronbach's alpha for the was .79 for the car version and .81 for the dealer sales version and .73 for the dealer after-sales version.

Validity of the scales was established in an earlier study (Bloemer, 1993). In this study content, construct and discriminant validity was extensively examined. From that study it can be concluded that the scale developed there and used here measuring commitment is valid.

Send correspondence regarding this article to:

Josee M. M. Bloemer
Limburg University Center
Faculty of Applied Economics
University Campus, Building D
B-3590 Diepenbeek, BELGIUM

THE EFFECTS OF ORGANIZATIONAL COMPLAINT RESPONSES ON CONSUMER SATISFACTION, WORD OF MOUTH ACTIVITY AND REPURCHASE INTENTIONS

Moshe Davidow, University of Haifa
James H. Leigh, Texas A&M University

ABSTRACT

This research investigates the link between organizational responses and post complaint consumer behavior, including satisfaction, word of mouth and intentions to repurchase. A conceptual model is introduced and empirically tested using structural equations modeling. The results are largely supportive of the hypothesized model. Satisfaction with the complaint response has a strong direct impact on word of mouth and intention to repurchase, however the relationship of word of mouth with intentions is not significant. The numbers of mail and telephone contacts made with the company to achieve problem resolution are negatively related to satisfaction, but not significantly related to subsequent word of mouth activity. Limitations of the study are discussed and directions for future research assessed.

BACKGROUND

Consumer complaining behavior is a topic of interest in marketing (Andreasen 1988). Such behavior is associated with a number of factors, including the consumer's personality (Richins 1982), attributions of blame (Folkes 1984), significance of the consumption event (Day 1984), probability of complaint success (Richins 1983, Day 1984), environmental influences (Singh and Wilkes 1991), and product importance (Richins 1985). In that a dissatisfied consumer may cease purchasing from a company, some organizations have even started to practice defensive marketing in an attempt to retain complaining customers (Fornell and Wernerfelt 1987). Understanding consumer complaining is thus an important area of inquiry.

Often-overlooked aspects of consumer complaint behavior are the organizational response to the complaint and its impact on a consumer's satisfaction, intentions to repurchase and word of mouth communications with others. Indeed, it is difficult to separate the concept of consumer

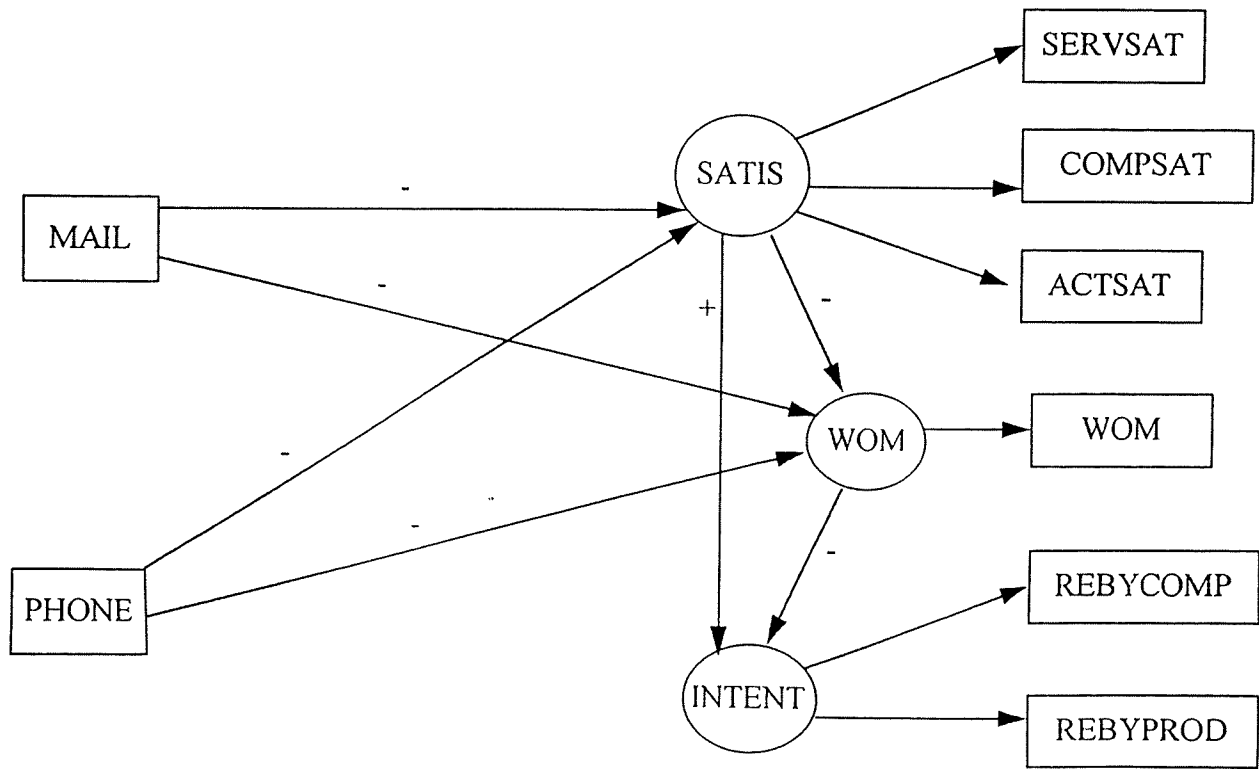
dissatisfaction from the concept of organizational responses to consumer complaints (Gilly 1987; Garrett, Meyers and Casey 1991), and it is misleading to study such behavior in isolation. Relatively little research has investigated this relationship (Blodgett, Granbois and Walters 1993; Clark, Kaminski and Rink 1992; Conlon and Murray 1996; Gilly 1987), and even this research has focused more on satisfaction with the response than on its outcomes.

The dimensions of organizational response to complaints have received some conceptual attention (see Davidow 1995), however they await full empirical testing. Gilly (1987) focused on speed of the response and the compensation as driving the complainant's satisfaction, but she did not examine the effect of satisfaction with the company's response on word of mouth. However, Blodgett, Granbois and Walters (1993) did find in their investigation that perceived justice has a positive effect on repurchase intentions, and a negative one on word of mouth; satisfaction was not explicitly examined. Based on Goodwin and Ross (1989), compensation and the personal interaction or service have strong impacts on satisfaction and repurchase intentions.

Satisfaction with the complaint response does positively affect repurchase intentions (Kolodinsky 1992), but it does not appear to dominate the repurchase decision and cannot restore repurchase intentions to that of a satisfied non-complainer (Halstead and Page 1992). The actions taken by the company do have a positive effect on repurchase intentions and a negative effect on word of mouth (Morris 1988), but the interaction of the complainant and the company's complaint handling can either enhance or detract from their satisfaction (Garrett et al. 1991).

It seems clear that research is needed that examines complaint relationships in a comprehensive manner. We present results of such research, that also extends previous work by examining complaint behavior in Israel, which has not been the focus of investigation heretofore.

Figure 1
Hypothesized Model



Hypothesized directions are shown by +/-

CONCEPTUAL FRAMEWORK AND HYPOTHESES

Our proposed model is displayed in Figure 1. It focuses on consumer satisfaction with complaint handling by an organization, word of mouth activities and repurchase intentions based on that satisfaction, and on the number of contacts a customer had with the company before a solution is achieved. Each construct and the linkages among the constructs are discussed below.

Conceptualization of Satisfaction

Satisfaction has been conceptualized several different ways in the literature, which may have had an effect on the results observed. Gilly (1987) conceptualizes satisfaction using an overall measure and a corresponding disconfirmation measure. Bearden and Teel (1983) consider disconfirmation an antecedent of satisfaction, and

measure satisfaction using four items adapted from Oliver (1980). Spreng et al. (1996) specify satisfaction as having two antecedents - - attribute and information satisfaction, and Tax and Chandrashekar (1992) use satisfaction with the product and with the service. Building on Gilly's (1987) perceptions of the organizational response to the complaint, Blodgett et al. (1993) focus instead on perceived justice as being an antecedent of repurchase intentions and word of mouth behavior. Oliver and Swan (1989) found that fairness is more important than is disconfirmation in producing satisfaction.

It is our belief that overall satisfaction is aggregative, based on Cadotte, Woodruff and Jenkins (1987) and Westbrook (1982), who argue that satisfaction is probably not a single emotion, but multidimensional. As such, we measure satisfaction as being manifested by the complainant's satisfaction with the service provided by the company's complaint department,

with the overall compensation, with the compensation amount, and with the actions taken.

How the complaint is handled is almost as important as the actual solution (Goodwin and Ross 1989, Lewis 1983). This view suggests that a consumer is dependent not only on the outcome, but also on the way that outcome is reached. The complaint communication process must be easy and clear for consumers, and the company representatives must be considerate and helpful. Garrett et al. (1991) found that this communication between the consumer and the organization is a key construct in most complaint management situations.

Consumer satisfaction with the overall compensation process rests on the company's ability to return the customer to the pre-complaint position. The compensation must take into account all of the consumer's financial and psychological costs (Andreassen 1988), including any losses incurred during the use of the product. Gilly (1987) found that repayment of any financial loss is positively related to consumer satisfaction. It seems reasonable then that a failure to be reimbursed for all costs or expenses related to the dissatisfaction will limit satisfaction recovery.

An organization's actions can also have an impact on a consumer's overall satisfaction with the complaint handling process. Consumers want a full explanation of why a problem occurred, and what actions will be taken by the organization to prevent a recurrence (Morris 1988). Lewis (1983) found that the actions taken to correct a problem are highly related to whether or not a consumer would repurchase. Consumers expect the company to provide a positive response, and will evaluate a company by its actions rather than its words alone.

Satisfaction and Word Of Mouth

Satisfaction with the organizational response to a complaint tends to be negatively related to a subsequent consumer's word of mouth activity (Bearden and Oliver 1985; Oliver and Swan 1989; TARP 1986). Moreover, there is a negative link between the perceived justice or fairness of the organizational complaint response and the likelihood of engaging in negative word of mouth behavior (Blodgett et al. 1993; Oliver and Swan 1989). In a study conducted in the Netherlands,

Richins (1987) reports that one way to control negative word of mouth is to encourage consumers to seek redress when dissatisfied and then handle these complaints to the consumer's satisfaction. Based on these findings, the following hypothesis is proposed:

H1: The more satisfied a consumer is with the complaint handling, the less likely the consumer is to engage in word of mouth activity.

Satisfaction and Repurchase Intentions

Consumer satisfaction is not an end in and of itself, but rather a means to the desired outcome of repurchase. If an organization does not give a consumer a good reason to come back, then that consumer will most likely transfer their allegiance to another brand. Complainants who perceive that justice has been served report higher repatronage intentions than those who perceive a lack of justice (Blodgett et al. 1993). There is a direct positive link between satisfaction and intentions to repurchase (Halstead and Page 1992; Kolodinsky 1992; TARP 1986). In light of these consistent findings, it is expected that:

H2: The more satisfied a consumer is with the organizational complaint handling, the higher the consumer's intention to repurchase the product.

Word Of Mouth and Intentions to Repurchase

Tax and Chandrashekar (1992) found that negative word of mouth has a significant impact on repurchase intentions. Based on Self Perception Theory (Bem 1972), if the consumer were to still buy the product after engaging in negative word of mouth, there would be an inconsistency between their words and actions. Having engaged in negative word of mouth activity, the consumer may find it difficult psychologically to repurchase the product. According to equity theory, victims are not hesitant to get even with those who treat them unjustly by retaliating against them (Walster, Walster and Berscheid 1978). Negative word of mouth and personal boycott are possible forms of retaliation against a company. Therefore, the

more a person engages in negative word of mouth, the more likely they are to not repurchase from that company. Considering these positions, it is expected that:

H3: The more a consumer engages in word of mouth, the less likely the consumer is to repurchase the product.

Antecedents to Satisfaction and Word of Mouth

Consumers evaluate the costs and benefits of complaining (Andreasen 1988, Richins 1980), and anything that raises the cost of complaining is going to lower the overall utility or satisfaction from the complaint response, all else equal. The number of times a consumer has to contact the company by mail or by phone to solve the problem is indicative of the effort expended and would serve to reflect the cost to the consumer. Extra contacts could be the result of missed communications, complications in the recovery process, high expectations, or some other factor. It seems likely that the more consumers feel the process is dragging out, the more effort they will likely expend in writing or calling the company and the less satisfied they will be with whatever outcome they are able to obtain. Based on these ideas, it is expected that:

H4a: The more mail contact a consumer has with an organization to solve a problem, the less satisfied the consumer will be with the outcome.

H4b: The more phone contact a consumer has with an organization to solve a problem, the less satisfied the consumer will be with the outcome.

Consumer complaining behavior is a complex, dynamic process, where negative word of mouth is primarily dependent upon the consumers post-complaint perception of justice (Blodgett et al. 1993). Each time a consumer contacts the company by mail or phone their perceived overall cost of complaining increases, thus lowering their satisfaction level and perceived justice redress. Consumers who perceive little likelihood of successful redress are likely to engage in word of

mouth behavior (Blodgett et al. 1993). Therefore, the more times a consumer has to write or call the company to achieve resolution to their problem, the more likely is their propensity for engaging in word of mouth behavior with others. In light of these findings, it is expected that:

H5a: The more mail contact a consumer has with an organization to solve a problem, the more likely the consumer is to engage in word of mouth activity.

H5b: The more phone contact a consumer has with an organization to solve a problem, the more likely the consumer is to engage in word of mouth activity.

METHODOLOGY

Sample

Data were gathered in Israel from 775 respondents to a questionnaire on customer complaint handling satisfaction. A self-administered questionnaire and postage-paid return envelope were mailed by an Israeli consumer goods manufacturer to those customers who had voiced a complaint regarding an inexpensive (<\$5) product. Questionnaires were mailed approximately two months after the complaint was handled. No callbacks were made. The total mailing was 1,513, of which 23 were returned due to incorrect addresses, yielding an effective response rate of 52.0%. Due to partial missing data, 109 responses were not included in the analysis, leaving a total of 666 usable surveys, or 44% of the total number mailed.

Measures Used

The selections of items for our measures are based on conceptual definitions. The operationalization of these variables is detailed in the Appendix.

Satisfaction Construct. Satisfaction is measured by four items: overall satisfaction with the service received from the complaint handling department (SERVSAT), satisfaction with the action taken by the company to resolve the

Table 1
Correlations and Variance/Covariance

		1	2	3	4	5	6	7	8	9	Mean	S.D.
SERVSAT	[1]	1.10	0.49	0.72	0.68	-0.14	0.33	0.36	-0.10	-0.21	4.06	1.05
AMTSAT	[2]	0.32	0.38	0.69	0.48	-0.14	0.22	0.28	-0.08	-0.15	1.71	0.62
COMPSAT	[3]	0.91	0.51	1.45	0.71	-0.09	0.31	0.36	-0.12	-0.21	3.57	1.20
ACTSAT	[4]	0.63	0.26	0.75	0.77	-0.10	0.31	0.39	-0.12	-0.23	4.31	0.88
W O M	[5]	-0.18	-0.11	-0.13	-0.11	1.50	-0.13	-0.09	0.09	-0.04	2.06	1.22
REBYCOMP	[6]	0.20	0.08	0.21	0.15	-0.09	0.32	0.40	-0.07	-0.07	4.78	0.57
REBYPROD	[7]	0.42	0.19	0.48	0.38	-0.13	0.25	1.24	0.00	-0.11	4.16	1.11
MAIL	[8]	-0.05	-0.02	-0.08	-0.06	0.06	-0.02	0.00	0.27	-0.32	1.03	0.52
PHONE	[9]	-0.12	-0.05	-0.14	-0.11	-0.03	-0.02	-0.07	-0.09	0.29	0.24	0.54

Note: Correlations are in bold, above the diagonal
 Variances are not bolded, on the diagonal
 Covariances are not bolded, below the diagonal

complaint (ACTSAT), satisfaction with the compensation amount (AMTSAT), and satisfaction with the overall compensation (COMPSAT). Due to the high degree of intercorrelation between AMTSAT and COMPSAT, the satisfaction construct is represented two different ways by using one of these variables in the formulation. Using the AMTSAT item, the three satisfaction items have a coefficient alpha of .79, whereas using COMPSAT, the satisfaction items have a coefficient alpha of .88. Both construct operationalizations exhibit internal consistency (Nunnally and Bernstein 1994).

Other Constructs. Word of mouth is measured by a single-item scale. The Repurchase Intentions construct is represented by two items - intentions to repurchase from the company (REBYCOMP) and to repurchase the brand (REBYPROD). Coefficient alpha for the two intentions items is 0.56, which is somewhat low, but still acceptable (see Nunnally and Bernstein 1994). The two exogenous antecedents - the reported number of mail (MAIL) and phone (PHONE) contacts made by the customer - are measured as separate manifest variables.

RESULTS

Measurement Model Assessment

Evidence of convergent and discriminant

validity for the satisfaction and repurchase intentions constructs can be assessed by confirmatory factor analysis (Joreskog and Sorbom 1989). We used the measurement model to show that the two constructs are in fact separate. The correlations among the indicators and the variances and covariances are given in Table 1.

The two factor model using AMTSAT as one of the Satisfaction indicators fits the covariances closely ($c^2=2.89$, $df=4$, $p=.58$), in spite of the somewhat low indicator reliabilities for AMTSAT (.35), REBYCOMP (.33), and REBYPROD (.47). The associated unidimensional model c^2 value of 48.06 ($df=5$, $p=.00$) is significantly larger than the two-factor model ($Dc^2=45.17$, $df=1$, $p=.00$), indicating that the Satisfaction and Repurchase Intentions constructs can be considered to be separate and distinct. Further evidence is provided by the 8.4% improvement in the total coefficient of determination (TCD) of .908, the 8.3% improvement in the Adjusted Goodness of Fit Index (AGFI) of .994, and the large reduction of 84.6% in the Root Mean Square Residual (RMSR) to .008 for the two-factor over the unidimensional construct model. The Phi coefficient of .67 ($t=15.18$, $p<.001$), indicates that the two constructs are related, but non-overlapping. For this analysis, significant ($p<.05$) t-statistics were observed for all estimated parameters (all $t's>9$), and the largest standardized residual of only 1.06 provides support for the model's adequacy, based on the guidelines offered by Bagozzi and Yi

(1988).

For the second confirmatory analysis in which COMPSAT was used in place of AMTSAT as a satisfaction indicator, a somewhat-larger c^2 of 5.21 ($df=4$, $p=.266$) was observed, a smaller AGFI (.988), a larger RMSR (.011), but with a larger overall TCD (.935). As was true for the confirmatory analysis with AMTSAT as a satisfaction indicator, all estimated parameters are significant (all t -values > 9), and the largest standardized residual of 1.8 provides support for the adequacy of the model. Compared to the unidimensional construct model, a reduction in c^2 of 50.64 ($df=1$, $p=.000$), an improvement of 5.9% in TCD and of 9.3% in AGFI, and a reduction of 81.0% in RMSR provide further evidence that the satisfaction and repurchase intentions constructs are distinct. The Phi coefficient of the correlation among the two constructs is .645 ($t=14.99$, $p<.001$), indicating that the two constructs are related, but not overlapping.

An analysis of the parameter estimates for the construct indicators, the associated squared multiple correlations, and the average squared variance can provide indirect evidence of the convergent validity of the satisfaction and repurchase intentions constructs. With only two indicators of repurchase intentions and three of satisfaction, the ability to assess convergent validity is limited to an analysis of the descriptive characteristics of each construct. With the exception of AMTSAT, the squared multiple correlations (all above .675) and the factor loadings (.822 - .858) for the satisfaction construct indicators are indicative of indicator reliability and convergent validity of the construct, respectively (Bagozzi and Yi 1988). AMTSAT is somewhat low on indicator reliability (.35) and its loading of .592 on the latent construct is lower than the other construct indicators; rather than eliminating this variable from the set of satisfaction indicators, the decision was made to retain this somewhat-fallible measure and use it in structural analyses in place of COMPSAT as one of the indicators. For the analysis with AMTSAT, the average squared variance is .57, whereas it is .70 when COMPSAT is included. Based on these results, the satisfaction construct appears to demonstrate reasonable convergent validity.

In the case of repurchase intentions, the squared multiple correlations for the two indicators are somewhat low, ranging from .33 to .47, but the factor loadings are fairly reflective of convergence (.58 to .68), based on Bagozzi and Yi (1988). With an average squared variance of .40, which is below the recommended .50 given by Bagozzi and Yi, it must be concluded that there is only limited evidence of convergent validity for the Repurchase Intentions construct.

Structural Model Testing Approach

The analysis procedure used to test the specified hypotheses is structural equations modeling with latent variables using the LISREL 7 program (Joreskog and Sorbom 1989). The covariance matrix used to analyze the data is reproduced in the lower diagonal of Table 1.

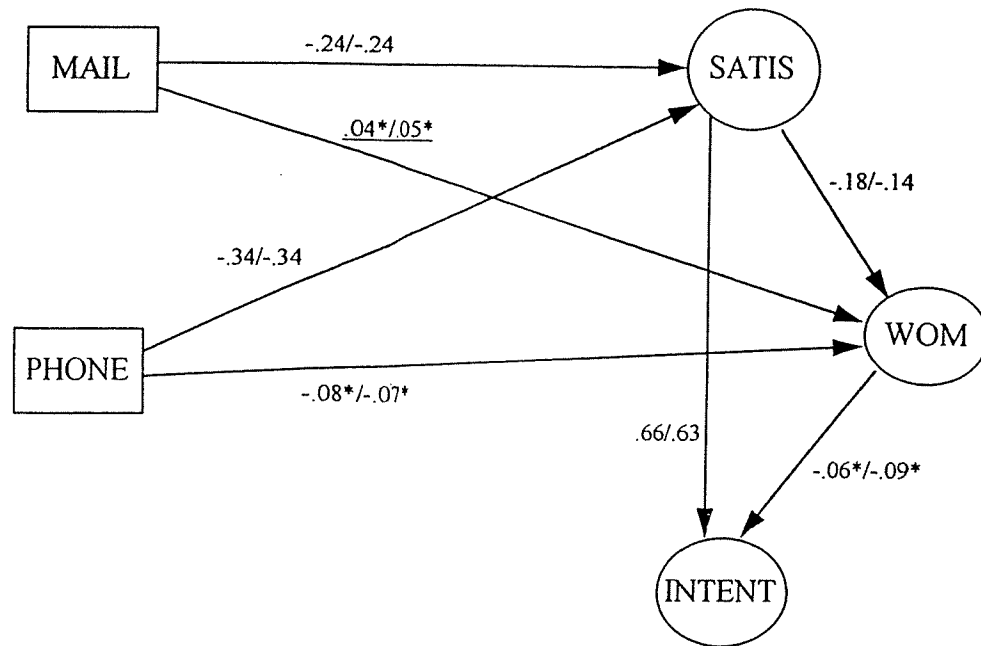
The Satisfaction and Repurchase Intentions endogenous constructs are specified as unobserved latent variables, with three and two indicators, respectively. In the case of Satisfaction, SERVSAT, ACTSAT, and either AMTSAT or COMPSAT form the indicator set. The SERVSAT parameter was used as the reference indicator to fix the scaling for the Satisfaction construct. REBYCOMP and REBYPROD are used to reflect the Repurchase Intentions construct, and the REBYCOMP parameter was fixed at 1.00 to set the scaling. Word of Mouth is represented by a single item, specified as measured without error. The two exogenous antecedents, MAIL and PHONE, are specified as directly-observed, fixed-X indicators, and assumed to be measured without error and uncorrelated with one another.

For corroborative purposes, a variable was also constructed as the sum of MAIL and PHONE, and specified as the only fixed-X antecedent. As the results were comparable to the others, attention will given here to the results of analyses based on MAIL and PHONE.

Structural Model Analysis Results

Preliminary Analyses. As can be seen in Figure 1, the only relationships for which a priori hypotheses were not specified for testing involve the direct effects of the MAIL and PHONE antecedents on Repurchase Intentions. An initial

Figure 2
Estimated Model with Standardized Path Coefficients



All Coefficients are standardized.

Asterisks indicate relationships hypothesized, but not significant at .05.

Model results are with the AMTSAT variable before the slash (/), and the COMPSAT variable after the slash. Underlined coefficients not only were not significant, but were not in the hypothesized direction.

set of analyses was completed that included these links as well as the ones hypothesized. Irrespective of whether AMTSAT or COMPSAT were used as one of the Satisfaction construct indicators, the inclusion of the two additional parameters did not significantly improve the fit over the *a priori* hypothesized model ($Dc2=2.60$ and 2.01 , respectively, $df=2$, $p > .25$), and neither of the two parameter estimates were significant (both t 's < 1.5). Based on these results, it can be concluded that key direct structural relationships were not omitted that would undermine the model estimation process.

Analyses of Substantive Relationships. The key results of the LISREL analyses of the predicted structural relations are depicted in Figure 2 for the models using either AMTSAT or COMPSAT as one of the respective Satisfaction construct indicators. This figure provides the

standardized structural relationships among the endogenous constructs and of the fixed-X antecedents with Satisfaction and Word of Mouth; more detailed information regarding specific construct and error parameter estimates as well as model diagnostic information are given in Table 2. Attention will first be given to discussing the fit results of the analyses before directing primary attention to assessing the extent to which results conform with the hypotheses framed.

As can be seen in Table 2, the $c2$ for the model based on AMTSAT as a Satisfaction construct indicator is 17.07 ($df=15$, $p=0.315$), and for the one based on COMPSAT as an indicator is 19.26 ($df=15$, $p=.202$). The respective AGFI's are $.985$ and $.983$, and the RMSR's are fairly low at $.013$ and $.014$. The TCD's of $.128$ and $.126$, respectively, are not very large, but the models do fit closely the observed covariances. The largest respective standardized

Table 2
Model Results

	<u>Parameter</u> <u>Loadings</u>	<u>Standard</u> <u>Error</u>	<u>Error</u> <u>Terms</u>	<u>Standard</u> <u>Error</u>
Measurement Model				
Satisfaction Construct			.65/.67	.06/.05
SERVSAT	1.00*/1.00*	.00/.00	.36/.33	.04/.03
AMTSAT/COMPSAT	.43/1.18	.03/.05	.25/.38	.02/.04
ACTSAT	.85/.83	.04/.04	.24/.25	.03/.02
Word Of Mouth Construct			1.44/1.46	.08/.08
WORD OF MOUTH	1.00*/1.00*	.00/.00	.00/.00	.00/.00
Repurchase Intentions Construct			.06/.06	.01/.01
REBYCOMP	1.00*/1.00*	.00/.00	.21/.21	.02/.02
REBYPROD	2.28/2.24	.25/.25	.67/.68	.07/.07
Structural Model				
SatisfactionYWOM (β_{21})	-.25/-.20	.07/.06		
SatisfactionYIntent (β_{31})	.25/.24	.03/.03		
WOMYIntent (β_{32})	-.02/-.02	.01/.01		
Antecedents-Satisfaction				
MailYSatisfaction (γ_{11})	-.39/-.40	.07/.07		
PhoneYSatisfaction (γ_{12})	-.54/-.54	.07/.07		
Antecedents-WOM				
MailYWOM (γ_{21})	.09/.11	.10/.10		
PhoneYWOM (γ_{22})	-.17/-.15	.10/.10		
Summary Statistics				
χ^2	17.07/19.26			
Degrees of Freedom	15/15			
p	.315/.202			
AGFI	.985/.983			
RMSR	.013/.014			
TCD	.128/.126			

Note: Model results are with the AMTSAT variable before the slash (/) and the COMPSAT variable after the slash.

— = Underline denotes t-value not significant at $p = .05$

* = Fixed at 1.0 for scaling purposes

residuals are 1.85 and 2.00, and the Q-plots are linear, providing further evidence of a reasonable fit (Bagozzi and Yi 1988). In each model test, all parameter estimates were found to be statistically significant, except for the relationships of the MAIL and PHONE antecedents with Word of Mouth and the relationship of Word of Mouth with

Repurchase Intentions. With the exception of H5b, which predicted a positive relationship between the reported number of phone calls made to the company and word of mouth behavior, in every other instance for which an hypothesis were posited, the direction of the relationship is congruent with prediction.

In terms of substantive relationships, the standardized parameter estimates shown in Figure 2 provide a useful indication regarding their relative importance. This information will be used in conjunction with the parameter estimate results given in Table 2 and associated t-values to evaluate each hypothesis. As shown in Figure 2, the relationship of Satisfaction with Repurchase Intentions is positive and significant (.66 and .63, t 's=9.24 and 9.44), which provides strong support for H1. It is evident that the more satisfied a consumer is with the complaint handling received, the more likely is their repurchase intention. The relationship of Satisfaction with Word of Mouth (-.18 and -.14, t 's = -3.83 and -3.27) is negative and significant, providing support for H2. By satisfying the complainant, not only does the company have a good chance of retaining their business, but they also reduce the chance of negative word of mouth.

The relationship of Word of Mouth with Repurchase Intentions (-.06 and -.09, t 's = -1.38 and -1.85, respectively) is in the hypothesized direction, and approaches significance in the case of the second model which used COMPSAT as one of the Satisfaction construct indicators. However, neither of these coefficients for the two models indicate that the WOM-Intentions relationship is a strong one, providing very weak support for H3, at best. Nested model tests of the effect of the direct path between Word of Mouth and Repurchase Intentions indicate that the fit of the two models are not significantly different if AMTSAT is used as a Satisfaction construct indicator ($Dc2=0.88$, $df=1$, $p>.25$), or if COMPSAT is used ($Dc2=3.41$, $df=1$, $p>.05$). One possible reason for the nonsupport for H3 may be attributed to the manner by which word of mouth was measured; the valence of the word of mouth was not recorded, only the amount. In retrospect, valence should have been assessed as well. This was a methodological oversight which unfortunately was not corrected in time. Had it been possible to separate reports of positive and negative word of mouth, more specific directional hypotheses could possibly have been formulated, and the results might have revealed a Word of Mouth - Repurchase Intentions linkage.

Based on H4a and H4b, negative relationships are predicted regarding Satisfaction and the

number of times the customer had to contact the company by mail (MAIL) or by telephone (PHONE), respectively, to receive a response. The results are fully supportive of these hypotheses, with the results for MAIL (-.39 and -.40, t 's = -5.52 and -5.79) and PHONE (-.54 and -.54, t 's = -7.80 and -7.99), indicating that the greater the effort taken by a complaining customer to receive a response, the less satisfied is the customer with aspects surrounding the outcome.

In contrast, H5a and H5b predict a positive relationship between mail and phone contacts with a company and subsequent word of mouth activity. The results are not supportive of these hypotheses. In the case of the MAIL-Word of Mouth relationship, the relationship is positive as predicted, but non-significant (.09 and .11, t 's=0.94 and 1.10). However, in the case of PHONE -Word of Mouth, the relationship is negative and non-significant (-.17 and -.15, t 's = -1.76 and -1.51, p 's > .05). The underlying basis for this negative relationship is not clear. Additional research will be necessary to obtain a more complete understanding of the basis, magnitude and stability of this negative relationship. Nested model tests of the value of including MAIL- and PHONE-Word of Mouth relationships in the hypothesized model reveal that if AMTSAT is used as one of the Satisfaction construct indicators, there is value in including the linkages ($Dc2=6.09$, $df=2$, $p<.05$), and if COMPSAT is instead used, the inclusion of the two parameters approaches significance ($Dc2=5.50$, $df=2$, $p<.07$). These results are due in part to the strength of the negative PHONE-Word of Mouth relationships, but they are also partly due to the indirect effects of MAIL and PHONE on Repurchase Intentions that operated through Word of Mouth.

DISCUSSION

Overall, the results provide much support for the hypothesized model, helping to clarify the relationships and showing the importance of complainant contacts with the organization to their satisfaction with the complaint handling. Specifically, satisfaction has strong direct impacts on word of mouth and intentions to repurchase, but the indirect effect of satisfaction on repurchase

intentions operating through word of mouth is weak, at best. The examined antecedents to satisfaction of the numbers of mail and phone contacts made by a complainant have strong direct relationships with satisfaction; the relationships of the antecedents with word of mouth and repurchase intentions are indirect, operating through satisfaction. This study replicates and extends previous consumer complaint research conducted in the United States by our performing this work in Israel, a Middle Eastern country recently upgraded from developing country to industrial country status, thus increasing the generalizability of these findings.

This study is not without limitations, which should be mentioned before addressing the managerial implications. The survey instrument included only one question that pertained to the complainer's reported word of mouth activity; additional questions regarding the nature and direction of the word of mouth are needed in future research to gain a more complete understanding of the model relationships. In this way, both the limitation of using a single item to reflect a behavioral domain and the limitation of not differentiating between positive and negative word of mouth would be addressed. Similarly, the antecedent measures of the number of mail and phone contacts made by the complainer are based on respondent self-reports, and are subject to the same errors and biases characteristic of self-report data.

There are several managerial implications that are indicated by our research. First, we have found a significant negative relationship between the number of contacts made by the consumer to the organization in the hopes of getting a suitable response, and their satisfaction with that response. This finding extends Gilly's (1987) finding that response speed is one of the factors driving consumer satisfaction with the recovery; not only is speed important, but also is the effort expended. The implication for the organization is to make their policies more customer friendly in order to minimize the number of contacts that need to be made. It would seem plausible that these results would hold true also in cases of face to face interactions between customers and complaint handlers, perhaps in a retail setting. This would be an interesting area for future research.

Second, satisfaction can lessen the amount of negative word of mouth. Given that negative word of mouth reaches twice as many people as positive word of mouth (TARP 1986), any method that can decrease the negative word of mouth is viable. In this case, increasing customer satisfaction with the complaint response will significantly decrease the negative word of mouth in the marketplace, emphasizing Richin's (1987) finding that one way to control negative word of mouth is to handle one's complaints effectively.

Third, satisfaction with the complaint response is shown here to have a very strong impact on the intention to repurchase, which replicates previous research (Blodgett, Granbois and Walters 1993; Goodwin and Ross 1989). In this respect, there appears to be no marked difference between consumers in the United States, the Netherlands and Israel, leading us to expect that the relationship is generalizable to other populations as well. It seems reasonable to conclude that proper complaint management can significantly add to a company's bottom line.

Finally, research generally has focused on how word of mouth affects other non-involved consumers and not on whether it impacts future purchase intentions of the person providing the word of mouth. The lack of a significant effect on repurchase intentions implies that it is not necessarily the case that word of mouth by a complainant impacts their intention to repurchase; however, the lack of a delineation between positive and negative word of mouth in our study precluded a more comprehensive examination of the relationship. Future research should definitely look at valance of the word of mouth as a possible link in the relationship. There are also other possible explanations for the lack of a significant result. Perhaps there is a moderating variable that we failed to capture, or perhaps there is an attribution issue (internal or external) at work. It is hoped that further research on this and other consumer complaint behavior relationships will continue in the future.

REFERENCES

- Andreasen, Alan R. (1988), "Consumer Complaints and Redress: What We Know and What We Don't Know," *The Frontier of Research in the Consumer Interest*, E. Scott Maynes and the ACCI Research Committee,

- (eds.), American Council on Consumer Interests, 675-722.
- Bagozzi, Richard P. and Youjae Yi (1988), "On The Evaluation of Structural Equation Models," *Journal of the Academy of Marketing Science*, 16, (Spring), 74-94.
- Bearden, William O. and Jesse E. Teel (1983), "Selected Determinants of Consumer Satisfaction and Complaint Reports," *Journal of Marketing Research*, 20, (February), 21-28.
- Bearden, William O. and Richard L. Oliver (1985), "The Role of Public and Private Complaining in Satisfaction With Problem Resolution," *Journal of Consumer Affairs*, 19, (2), 222-240.
- Bem, Daryl J. (1972), "Self Perception Theory," *Advances in Experimental Social Psychology*, Leonard Berkovitz (ed.), 6, New York: Academic Press, 1-62.
- Blodgett, Jeffrey G., Donald H. Granbois and Rockney G. Walters (1993), "Toward An Integrated Conceptual Model of Consumer Complaining Behavior," *Journal of Retailing*, 69, (Winter), 399-428.
- Cadotte, Ernest R., Robert B. Woodruff and Roger L. Jenkins (1987), "Expectations and Norms in Models of Consumer Satisfaction," *Journal of Marketing Research*, 24, (August), 305-314.
- Clark, Gary L., Peter F. Kaminski and David R. Rink (1992), "Consumer Complaints: Advice On How Companies Should Respond Based On An Empirical Study," *Journal of Consumer Marketing*, 9, (Summer), 5-14.
- Conlon, Donald E. and Noel M. Murray (1996), "Customer Perceptions of Corporate Responses to Product Complaints: The Role of Explanations," *Academy of Management Journal*, 39, (4), 1040-1056.
- Davidow, Moshe (1995), "Organizational Response to Consumer Complaints: Effective and Efficient Complaint Management," *Enhancing Knowledge Development in Marketing*, 6, Barbara Stern and George Zinkhan, (eds.), Chicago, IL: American Marketing Association, 344-345.
- Day, Ralph L. (1984), "Modeling Choices Among Alternative Responses to Dissatisfaction," *Advances in Consumer Research*, Provo, UT: Association for Consumer Research, Thomas C. Kinneer, (ed.), 496-499.
- Folkes, Valerie S. (1984), "Consumer Reactions to Product Failure: An Attributional Approach," *Journal of Consumer Research*, 10, (March), 398-409.
- Fornell, Claes and Birger Wernerfelt (1987), "Defensive Marketing Strategy by Customer Complaint Management: A Theoretical Analysis," *Journal of Marketing Research*, 24, (November), 337-346.
- Garrett, Dennis E., Renee A. Meyers and John Camey (1991), "Interactive Complaint Communication: A Theoretical Framework and Research Agenda," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 4, 62-79.
- Gilly, Mary C. (1987), "Postcomplaint Processes: From Organizational Response to Repurchase Behavior," *Journal of Consumer Affairs*, 21, (2), 293-313.
- Goodwin, Cathy and Ivan Ross (1989), "Salient Dimensions of Perceived Fairness in Resolution of Service Complaints," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 2, 87-92.
- Halstead, Diane and Thomas J. Page Jr. (1992), "The Effects of Satisfaction and Complaining Behavior on Consumer Repurchase Intentions," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 5, 1-11.
- Joreskog, Karl G. and Dag Sorbom (1989), *LISREL 7: A Guide to the Program and Applications*, Chicago, IL: SPSS, Inc.
- Kolodinsky, Jane (1992), "A System for Estimating Complaints, Complaint Resolution and Subsequent Purchases of Professional and Personal Services," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 5, 36-44.
- Lewis, Robert C. (1983), "Consumers Complain - What Happens When Business Responds?" *International Fare in Consumer Satisfaction and Complaining*, Ralph L. Day and H. Keith Hunt, (eds.), Bloomington, IN: Bureau of Business Research, 88-94.
- Morris, Susan V. (1988), "How many Lost Customers Have You Won Back Today? An Aggressive Approach to Complaint Handling in the Hotel Industry," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 1, 86-92.
- Nunnally, Jum C. and Ira H. Bernstein (1994), *Psychometric Theory*, 3rd Edition, New York, NY: McGraw-Hill.
- 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. 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.
- Richins, Marsha L. (1980), "Consumer Perspectives of Costs and Benefits Associated With Complaining," in *Refining Concepts and Measures of Consumer Satisfaction and Complaining Behavior*, H. Keith Hunt and Ralph L. Day, (eds.), Bloomington, IN: Bureau of Business Research, 50-53.
- Richins, Marsha L. (1982), "An Investigation of Consumer Attitudes Toward Complaining," *Advances in Consumer Research*, 9, Andrew Mitchell, (ed.), Provo, UT: Association for Consumer Research, 502-506.
- Richins, Marsha L. (1983), "Negative Word of Mouth by Dissatisfied Consumers: A Pilot Study," *Journal of Marketing*, 47, (Winter), 68-78.

- Richins, Marsha L. (1985), "The Role of Product Importance in Complaint Initiation," *Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, H. Keith Hunt and Ralph L. Day, (eds.), Bloomington, IN: Bureau of Business Research, 50-53.
- Richins, Marsha L. (1987), "A Multivariate Analysis of Responses to Dissatisfaction," *Journal of the Academy of Marketing Science*, 15, (Fall), 24-31.
- Singh, Jagdip and Robert E. Wilkes (1991), "A Theoretical Framework for Modeling Consumer's Response to Marketplace Dissatisfaction," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 4, 1-12.
- Spreng, Richard A., Scott B. MacKenzie and Richard W. Olshavsky (1996), "A Reexamination of the Determinants of Consumer Satisfaction," *Journal of Marketing*, 60, (July), 15-32.
- TARP (1986), "Consumer Complaint Handling in America: An Updated Study," *Office of Consumer Affairs*, Washington, D.C.: Technical Assistance Research Programs.
- Tax, Stephen S. and Murali Chandrashekar (1992), "Consumer Decision Making Following a Failed Service Encounter: A Pilot Study," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 5, 55-68.
- Walster, Elaine, G. William Walster and Ellen Berscheid (1978), *Equity Theory and Research*, Boston, MA: Allyn and Bacon.
- Westbrook, Robert A. (1982), "Consumer Satisfaction and the Phenomenology of Emotions During Automobile Ownership Experiences," *International Fare in Consumer Satisfaction and Complaining*, Ralph L. Day and H. Keith Hunt, (eds.), Bloomington, IN: Bureau of Business Research, 88-94.

experience? Coded 0, 1-3 = 1, 4-6 = 1, 7-10 = 3, 11+ = 4.

Repurchase Intentions

REBYPROD - Intention to repurchase this product, coded 1-5 anchored by definitely would repurchase (5), and definitely would not repurchase (1).

REBYCOMP - Intention to repurchase from this company, coded 1-5 anchored by definitely would repurchase (5), and definitely would not repurchase (1).

Antecedent Variables

MAIL - The number of times a customer contacted the company by mail.

PHONE - The number of times a customer contacted the company by phone.

Both variables coded 0, 1, 2+ = 2.

Send correspondence regarding this article to:

Moshe Davidow
Graduate School of Business
University of Haifa
Mount Carmel, Haifa 31905
ISRAEL

APPENDIX

Variable Operationalization

Satisfaction

SERVSAT - Overall satisfaction with the service received from the complaint department, coded on a scale of 5, anchored by very satisfied (5), and extremely dissatisfied (1).

ACTSAT - Satisfaction with the action taken by the company to resolve the complaint, coded 1-5 anchored by very satisfied (5), and extremely dissatisfied (1).

AMTSAT - Satisfaction with the compensation amount, coded 1-3 anchored by more than expected (3), and less than expected (1).

COMPSAT - Satisfaction with the overall compensation, coded 1-5 anchored by very satisfied (5), and extremely dissatisfied (1).

Word of Mouth - How many people did you tell about this

SATISFACTION WITH MATERIAL POSSESSIONS AND GENERAL WELL-BEING: THE ROLE OF MATERIALISM

M. Joseph Sirgy, Virginia Tech
Dong-Jin Lee, University of Western Australia
Val Larsen, Truman State University
Newell Wright, James Madison University

ABSTRACT

We theorize that overall life satisfaction is largely a composite of satisfaction from various life domains. The focus of this study is the material life domain, a psychological construct that groups consumer experiences related to the possession of economic goods. Consistent with our overarching hypothesis, we hypothesized that satisfaction with material possessions influences overall life satisfaction. We further hypothesized that the degree of influence material satisfaction has on life satisfaction is moderated by materialism, i.e., involvement with possessions. Specifically, material possessions will more strongly influence life satisfaction for materialistic than for non-materialistic individuals. We also hypothesized that satisfaction with material possessions is influenced by materialism. Specifically, those who are materialistic are more likely to be dissatisfied with their possessions because they are more likely to have high possession expectations. The model's constructs were operationalized in the context of a survey that was administered to about 300 college students. The data provided general support for the model.

INTRODUCTION

Quality-of-Life (QOL) has been conceptualized in various ways depending on the theoretical perspective used to define the construct (Sirgy, Meadow, and Samli 1995). From an economic perspective, QOL has been conceptualized in terms of income and utility. That is, many economists use accumulation of wealth, measured by income, as a surrogate measure of QOL. From a psychology perspective, QOL has been quite often conceptualized as life satisfaction (or subjective well-being). In this study, QOL is defined as overall life satisfaction.

In defining QOL in this way, this study builds upon the work of Campbell, Converse, and Rodgers (1976), who have suggested that life

satisfaction is a composite of satisfaction in various life domains (e.g., work, family, health, leisure etc). In other words, life satisfaction represents an overall attitude about one's life space and life domains within a life space (Campbell et al. 1976; Diener 1984; Brief et al. 1993). The greater the satisfaction within various life domains, the greater the satisfaction with life as a whole. The direct relationship between life domain specific satisfaction and overall life satisfaction is referred to as a "spillover" effect (Iris and Barrett 1972) or as "generalization" (Seeman, 1967). Turning to the specific focus of this study, the consumer life domain (Day 1987), people's assessment of quality of life corresponds closely to their acquisition and possession of economic resources (Douthitt, MacDonald, and Mullis 1992). And empirical research (Leelakulthanit, Day, and Walters 1991) suggests that satisfaction with material possessions significantly increases overall life satisfaction.

This study hypothesizes that materialism--emotional involvement in the possession life domain-- may lead to greater spillover from satisfaction with possessions to overall life satisfaction. We more specifically define materialism as the importance a consumer attaches to worldly possessions, in other words, as the belief or feeling that material objects are important and valuable (Belk 1984, 1985). Thus, materialistic people value material objects highly and devote more time, energy, and effort to the possession domain than others do. Consequently, they have fewer resources available in other life domains.

At present, little is known about the effects of materialism on the spillover of satisfaction with material possessions onto overall life satisfaction. The purpose of this study is, therefore, to explore this relationship. Marketers enhance overall life satisfaction (QOL) by providing material and non-material satisfactions. If the degree of spillover from possession satisfaction differs depending on the materialism of consumers, marketers can modify marketing programs (e.g., advertising) so

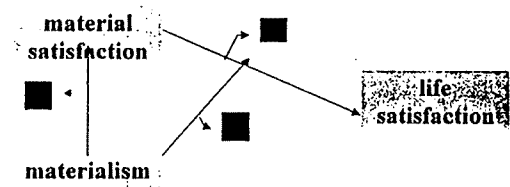
that they put more or less emphasis on material satisfactions depending on the degree of overall life satisfaction the emphasis is likely to produce. Advertising that emphasizes satisfaction with material possessions can be targeted at those with high materialism who are likely to derive satisfaction from a material emphasis. Conversely, advertising that features other life satisfactions may be targeted at people who are low in materialism and, therefore, are more likely to respond to and be pleased by products and images that are associated with other life domains.

This paper is organized as follows. We discuss first the spillover of possession satisfaction into overall life satisfaction, then the moderating effect of materialism on this spillover, and finally the relationship between materialism and possession satisfaction. Discussions on method, results, and implications follow.

A MODEL OF MATERIAL SATISFACTION, MATERIALISM, AND LIFE SATISFACTION

In the literature on quality of life (Campbell, Converse, and Rodgers 1976; Diener 1984; Scott and Stumpf 1984), the concept of *life satisfaction* is a well-accepted QOL indicator. We therefore focus our research on this concept. Our model of the relationship between key variables is represented in Figure 1, which shows that overall life satisfaction is mostly determined by satisfaction from the various life domains, including the material life domain. The material life domain groups an individual's experiences related to the possession of economic goods; thus, material satisfaction is the aggregate satisfaction produced by various material possessions. The influence of this material satisfaction on overall life is moderated by the extent of involvement in material life (materialism). Material satisfaction is likely to influence life satisfaction more strongly if an individual is materialistic. The model also shows that material satisfaction is influenced, negatively, by materialism, for materialistic people are likely to have high possessions expectations (Rudmin and Richins 1992; Wright and Larsen 1993), making them more likely to experience dissatisfaction with the possessions they actually have.

Figure 1
The Hypothesized Model



The Material Life Domain

Social psychologists have long recognized that the self is not a unidimensional construct. Self-concept is multidimensional because the mind of every person is constituted by multiple selves (Brewer and Nakamura 1984; Garza and Herringer 1987; Hoelter 1985; Markus 1977; McCall and Simmons 1978). These aspects of the self that form the self-concept are situated in various psychological *life domains*. Thus, people may have a self-concept in relation to education, family, health, job, friends, romantic relationships, etc. In other words, the psychological world is divided into life domains, and within each life domain, people have certain self-related, value laden beliefs that tend to be grouped, in their own minds, in life domains such as health, job, family, community, standard of living, and material possessions (Campbell, Converse, and Rodgers 1976; Burke and Tully 1977; Griffin, Chassin, and Young 1981; Scott and Stumpf 1984). The life domain that we focus on pertains to concerns related to economic goods (Douglas and Isherwood 1979; Kleine and Kernan 1991; Kleine, Kleine, and Kernan 1992, 1993; McCracken 1988; Morgan 1993). We refer to this life domain as the *material life domain*.

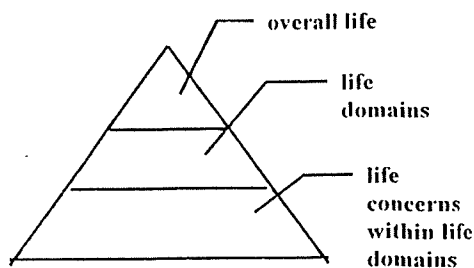
From the marketing perspective, Day (1978, 1987) and Leelakulthanit, Day, and Walters (1991) conceptualized the consumer life domain in terms of two dimensions: the acquisition and the possession of economic goods. *Acquisition of economic goods* refers to the domain of objects, persons, and events related to the purchase of goods that have economic value. Examples include the atmospherics, prices, assortment and quality of

goods, courtesy and helpfulness of personnel, and post-purchase service and warranties provided by local stores. The second dimension of the consumer life domain is the *possession of economic goods*. Possession of goods refers to the collection of objects that have monetary value (e.g., house/apartment, furniture, car/truck, clothing/accessories, and savings, etc.). In this paper we define the material life domain as involving only the *possession* of economic goods, not their acquisition.

Spillover Between Satisfaction with Material Possessions and Overall Life Satisfaction

To fully explain the relationship between satisfaction with the material possessions and overall satisfaction with life, we need to understand the concepts of vertical and horizontal spillover. *Spillover* between satisfaction with material possessions and overall life satisfaction can be represented using the hierarchy model shown in Figure 2 (Meadow 1988). The model is suggested by research in consumer satisfaction (e.g., Aiello, Czepiel, and Rosenberg 1977) and life satisfaction (e.g., Andrews and Withey 1976; Campbell, Converse, and Rodgers 1976) and is inferred from social gerontological research (e.g., Neugarten, Havighurst, and Tobin 1961). The model's basic premise is that life satisfaction is influenced by lower level concerns.

Figure 2
The Vertical Spillover



In this model, the affect within a life domain spills over *vertically* to the most superordinate domain (life in general), thus determining life

satisfaction. Most multiattribute attitude models use the same logic to explain and predict attitudes. For example, Fishbein and Ajzen (1975) have suggested that consumers attitudes toward a product, such as a car model or brand, are a direct function of their evaluations of the various attributes of the car as moderated by the belief strength associated with each attribute.

Satisfaction researchers have used this same logic to explain consumer satisfaction (e.g., Aiello, Czepiel, and Rosenberg 1977). In this case, the evaluation of each attribute expresses degree of satisfaction, and overall life satisfaction is held to be a function of satisfaction within each life domain (job, family, personal health, leisure, material possessions, and so forth). Satisfaction within a given life domain, in turn, is determined by satisfaction with the life conditions/concerns that make up that domain. For example, satisfaction with material possessions (the material life domain) may be determined by satisfaction with one's house, car, furniture, clothing, savings, jewelry, accessories, etc.

The extent to which satisfaction within a subdomain affects satisfaction of a superordinate domain in the hierarchy has been referred to in the quality-of-life literature as the *vertical spillover* (e.g., Sirgy, Hansen, and Littlefield 1994). It should be noted that spillover can be either bottom-up or top-down in vertical spillover. When satisfaction from a subordinate domain affects satisfaction in a superordinate domain, it is referred to as *bottom-up vertical spillover*. The spillover we have mentioned from satisfaction with material possessions to overall life satisfaction is an example of bottom-up vertical spillover. If the flow of affect is reversed and satisfaction in the superordinate domain spills over onto the subordinate domain, it is referred to as *top-down vertical spillover*. Empirical research has confirmed that top-down vertical spillover does occur (see Diener 1984 for a literature review). This concept suggests that life satisfaction may affect how people evaluate their material possessions. Those who are satisfied with life may be disposed to evaluate their possessions more positively than others do. Conversely, those who are dissatisfied with life may be disposed to evaluate their possessions more negatively.

Many QOL studies have empirically

demonstrated the bottom-up vertical spillover effect between satisfaction with specific life domains and overall life satisfaction (Diener 1984). With respect to the vertical spillover between satisfaction with material possessions and overall life, Dawson and Bamossy (1991) have shown that life satisfaction often plummets for people whose homes are destroyed by natural disasters. Leelakulthanit *et al.* (1991) have also demonstrated this effect using a consumer population in Thailand. They found that satisfaction with one's own acquisition and possession of material goods affected overall life satisfaction, especially for older and low income people. After examining much of the evidence on income and subjective well-being, Veenhoven (1991) concluded that income correlates highly and positively with subjective well-being. Applying these findings to our major constructs, this suggests that the material domain may affect overall life satisfaction. Hence, we hypothesize a bottom-up vertical spillover effect.

Turning to *horizontal spillover*, this concept refers to the influence affect in one life domain has on affect in another domain. Here, we are focusing on the spillover between life domains that are at the same hierarchical level in the overall hierarchy of life domains and concerns. For example, affect in the consumer domain may spill over into the family domain or job domain and visa versa. Although we acknowledge the effects of horizontal spillover between the consumer life domain and other domains, the model proposed in this paper does not deal with horizontal effects, only vertical spillover.

Based on the preceding discussion we introduce our first hypothesis.

Hypothesis 1: Overall life satisfaction is significantly influenced by affect or satisfaction in the material domain.

Involvement in a Life Domain Moderates the Domain's Effect on Overall Life Satisfaction

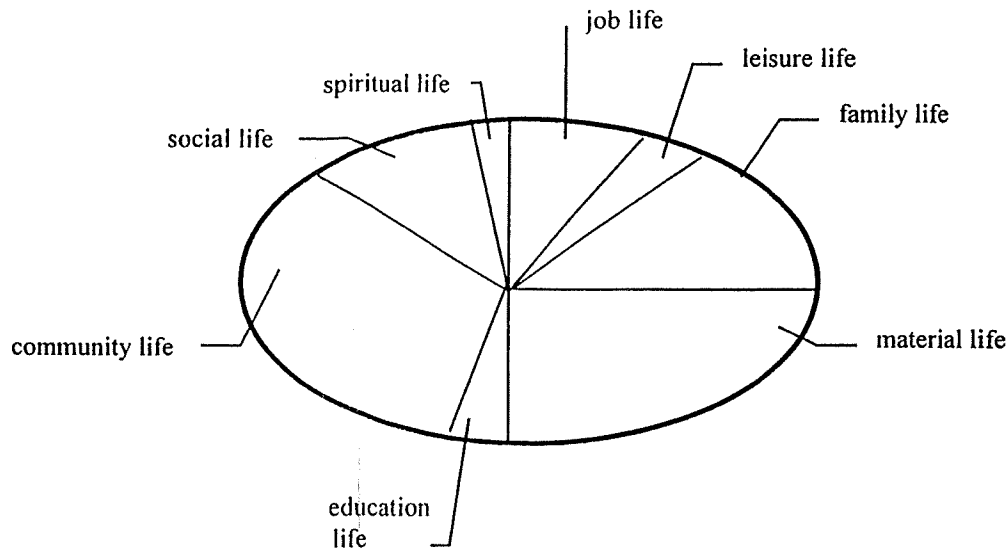
All life domains, including the consumer life domain and its subdomains, vary in their *salience*, i.e., some domains are more important than others. A number of prominent social psychologists (e.g., McCall and Simmons 1978;

Rosenberg 1978; Stryker 1968) have suggested that various identities (self-concepts reflective of certain life domains) are organized in hierarchies of salience that influence self-evaluations. For some people, their job may be the most important thing in life; for others, family may be most important; and, turning again to the focus of this study, for still other people, material possessions may be most important (see Figure 3). The concept of salience hierarchy involving life domains can illuminate our definition of materialism. For us, people are materialistic to the degree that the material domain (a sub-domain of the consumer life domain) is salient for them. Degree of salience can vary both absolutely and relatively, both within and between persons. For example, to a person suffering from depression, the material domain could be more salient than other domains (high relative salience) and yet still be less salient than it is for people who are mentally healthy (low absolute salience). Both the absolute and the relative levels of satisfaction are important. One's overall level of life satisfaction can be affected by changes in the salience of either the absolute or the relative salience of a life domain. But while it is important to keep in mind that materialism can vary absolutely between individuals and groups (a fact that is important in cross-ethnic and cross-national studies), in common parlance the term is most likely to be used to describe people within a group for whom material objects have high salience relative to other life domains that are important to the group.

Sociologists usually treat the concept of materialism in terms of people placing emphasis on things such as earning money and accumulating material possessions. For example, in a recent national study Easterlin and Crimmins (1988, 1991) argued that young people today are more materialistic than their counterparts in the 60's and 70's because they put more emphasis on earning a lot of money but less emphasis on work.

Turning to marketing, Belk and Pollay (1985) have defined materialism in relatively situational and concrete terms. In a study that focused on changes in the content of magazine advertising (and national materialism) over time, they defined materialism in terms of the degree to which ads emphasized appeals to pleasure and a comfortable or luxurious lifestyle. In the 80 years surveyed,

Figure 3
Various Life Domains



they found that materialism steadily increased. Offering a conceptual view of materialism that is more consistent with our theoretical perspective, other marketing researchers (e.g., Bloch and Richins 1983; Celsi and Olson 1988; Houston and Rothchild 1977) have distinguished between enduring and situational involvement with things. *Enduring involvement* reflects a general and permanent concern with a product class, while *situational involvement* reflects concern with specific situations or events related to specific purchases.

In the context of the material life domain (the world of economic goods and material possessions), we view materialism as enduring involvement in the material domain—a high level of cognitive and emotional involvement with material possessions. Our view of materialism as enduring involvement in the material life domain is consistent with Belk's view in other research (1983, 1984, 1985), wherein he defined materialism as interest in and concern for the ownership of things and the accumulation of possessions. "Materialism," he says, "reflects the importance a consumer attaches to worldly possessions. At the highest levels of materialism, possessions assume a central place in a person's life and are believed to provide the greatest sources of satisfaction and dissatisfaction in life"

(Belk 1984, p. 291). Viewed in this way, materialism is an individual difference factor related to the belief that material possessions are important in achieving happiness in life. Drawing upon these definitions of and this research on materialism, we offer the following hypothesis.

Hypothesis 2: The influence of material satisfaction on life satisfaction is stronger for more materialistic individuals than for less materialistic individuals.

The Effect of Enduring Involvement on Consumer Expectations and Dissatisfaction

Empirical work done to date seems to support a long-standing position of religious leaders and philosophers. In this stream of research, a consistent finding emerges: materialism is negatively related to overall life satisfaction. In a meta-analysis of studies treating the relationship (Belk 1985; Cole *et al.* 1992; Dawson and Bamossy 1990, 1991; Richins 1987; Richins and Dawson 1992), Wright and Larsen (1993) found a stable, medium-sized negative correlation. This negative relationship between materialism and life satisfaction can be explained as follows. Overall life satisfaction (QOL) is partly determined by satisfaction with material possessions. Satisfaction

with material possessions, in turn, is mostly determined by evaluations of one's actual state of material possessions compared to a set goal. *Materialists* tend to experience greater dissatisfaction with material possessions than nonmaterialists, which in turn spills over negatively upon overall life satisfaction. *Materialists* experience dissatisfaction with material possessions because they set material possession goals that are inflated and unrealistically high. This reasoning leads to the following hypothesis:

Hypothesis 3: Material satisfaction is negatively influenced by materialism: those who are more materialistic are likely to express greater dissatisfaction with their possessions than those who are less materialistic.

METHOD

Survey

In order to investigate the relationships illustrated in Figure 1 and test the hypotheses, a survey was used. The sample was drawn from undergraduate business students in eastern and mid-western universities. A total of 400 self-administered questionnaires were distributed in classes. Out of 450 questionnaires distributed, 293 generated usable data (response rate = 65 percent). The average age of respondents was 22.5. The sample was 57 percent male (167), 43 percent female (126). Each student received a questionnaire to fill out in the classroom. The questionnaire assured them that there were not right or wrong answers to the questions it posed. All that was required was their honest opinion.

Specific Measures

In this section we describe the specific measures used in the study.

Material Satisfaction. This construct was derived from the theoretical development of the model, and involved a multi-attribute composite index. This index reflects the satisfaction with ownership of specific material goods in a broad

range of product categories. This multi-attribute composite index was mathematically formulated as follows:

$$\text{OMS} = \text{Average SMi}$$

where OMS = overall material satisfaction
SMi = satisfaction with a specific category of material possessions (i)

The SMi component was measured as follows:

"If you own any of the following items, please indicate the extent to which you are satisfied/dissatisfied with possessing or owning them. Note that a person might like owning something—a classic car or a piece of property—even though they never use. Or they might be pleased both to own and to use the thing. On these items, indicate *only* how you feel about owning the item, *not* how you feel about using or consuming it. *Respond only to the items you own.*"

House or condominium

Consumer electronics (CD player, TV, VCR, computers, etc.)

Furniture and/or appliances

Private transportation (cars, trucks, motorcycles, and bicycles)

Clothing, accessories, and jewelry

Savings and investments

Each of these items was measured on a 7-point scale varying from "wonderful" (7) "good" (6), "satisfactory" (5), "neutral" (4), "unsatisfactory" (3), "bad" (2), "awful" (1), and "no opinion." The last category was treated as a missing value. Since all items represent different sub-constructs (not indicators of the same overall construct), no internal consistency-type of reliability was expected. The material satisfaction composite score for each subject was computed by averaging all six SMi scores.

The Materialism Measure. This construct was operationalized as the composite average score items selected from the following list:

I try to keep my life simple as far as material

possessions are concerned. (reverse coded)
 The things I own aren't all that important to me. (reverse coded)
 I put less emphasis on material things than most people I know. (reverse coded)
 I like a lot of luxury in my life.
 What I own or don't own has little effect on my personal happiness. (reverse coded)
 I often compare my standard of living with that of others.
 Where material possessions are concerned, generally I feel that more is better.
 It isn't important to own a nice car. (reverse coded)
 Owning property is or would be very satisfying for me.
 The things that money can buy are very important to me.
 I tend to view possessions as a burden, not as a blessing. (reverse coded)
 I like to have really nice things.
 Acquiring new material possessions is a low priority to me. (reverse coded)
 Owning material things is an important part of life.
 The more I have, the better I feel.
 I often indulge myself in buying things I don't really need.
 I keep my life simple by owning relatively few things.

Some of these items were adapted from Richins' materialism scales. We added other items. Responses to these items were recorded on a 7-point Likert-type scale. A single-factor confirmatory factor analysis was conducted and the results indicated an acceptable fit (Chi-Square = 71.22, $df = 14$, $p = .00$; GFI = .93; AGFI = .87; RMR = .07). Not all items were included in the composite average score, only those that loaded highly. The selected items were 4, 5, 8, 11, 12, 16, and 17, and the Cronbach alpha of these items was .6754.

The Life Satisfaction Measure. Life satisfaction was measured using a single-item Delighted-Terrible (D-T) scale. Subjects were asked "How do you feel about your life as a whole?" The D-T scale contained the following response categories: "delighted" (coded as 7),

"pleased" (6), "mostly satisfied" (5), "mixed feelings" (4), "mostly dissatisfied" (3), "unhappy" (2), and "terrible" (1).

The D-T measure is a well-established measure of subjective well being (Andrews and Withey 1976). The measure was reported to have a temporal reliability of .40 for a 6-months interval (Stock *et al.* 1982). Andrews and Withey (1976) reported high convergent validity with other self-report measures of life satisfaction and nomological validity, indicated by expected relationships with self-efficacy, marriage, and standard of living. Other positive and strong evidence of reliability and validity of the D-T measure was reported by Larsen, Emmons, and Diener (1983).

Satisfaction with Other Life Domains.

Satisfaction with other life domains such as job, family, finances, health, education, friends, leisure, neighborhood, and community was measured using the following single-item measures. Subjects were asked "How do you feel about the areas of your life that are listed below? Indicate whether you feel good or bad about each area of your life."

Your job situation
 Your family situation
 Your financial situation
 Your health
 Your education
 Your friends and associates
 Your leisure life
 Your neighborhood
 Your community
 Your spiritual life
 The taxes you pay
 Your environment (quality of air, water, land)
 Your political/economic freedom and independence
 Your housing situation
 Your cultural life
 Your social status

Responses to these items were measured using single-item Delighted-Terrible (D-T) scales, similar to the one used to measure life satisfaction. The D-T scale has been extensively used in quality-of-life research to measure domain-specific

Table 1
Lisrel Estimates

Sample	Dependent	Independent	Errorvar	R-Square
Entire (n=297)				
	LSG	<u>0.27</u> *POS (0.056) 4.77	.93 (0.076) 12.14	.072
	POS	-.00*IPOS (.058) -.13	.93 (.076) 12.14	.072
($\chi^2 = 8.45$, $df=1$, $p=0.003$, $GFI=.98$, $AGFI=.89$, $NFI=0.72$, $CFI=.73$, $RMR=.06$)				
High Involvement (n=132)				
	LSG	<u>.32</u> *POS (.083) 3.91	0.89 (.11) 8.06	.11
	POS	-.15*IPOS (.087) -1.76	0.98 (.12) 8.06	.023
($\chi^2 = 14.85$, $df=1$, $p=.00$, $GFI=.93$, $AGFI=.60$, $NFI=.54$, $CFI=.53$, $RMR=.12$)				
Low Involvement (n=151)				
	LSG	<u>.22</u> *POS (0.080) 2.77	.95 (.11) 8.63	.049
	POS	<u>.21</u> *IPOS (.080) 2.65	.95 (.11) 8.63	.049
($\chi^2 = .15$, $df=1$, $p=.90$, $GFI=1.0$, $AGFI=1.0$, $NFI=1.0$, $CFI=1.0$, $RMR=.0038$)				

1. LSG= Overall Life Satisfaction ; POS= Possession Satisfaction ; IPOS= Materialism
2. (1=low; 7=high)
3. significant coefficients: bold and underline

satisfaction as well as overall life satisfaction. (Andrews and Withey 1976; Larsen, Emmons, and Diener 1983; Stock *et al.* 1982).

RESULTS AND DISCUSSION

In this section, the study results are reported

by hypothesis. The spillover from material satisfaction to overall life satisfaction (H1), the moderating effect of materialism on the spillover relationship (H2), and the effect of materialism on material satisfaction are then discussed.

Table 2
Regression Analysis
(Dependent Variable = LSG: Overall Life Satisfaction)

Sample	Independent Variable	Pearson Correlation	Regression Estimate	p-value
Entire Group (n=297)				
	POS	.28	.17	.05
	LSD1	.21	.03	.38
	LSD2	.07	.03	.53
	LSD3	-.01	-.08	.13
	LSD4	.18	-.01	.81
	LSD5	.27	.11	.07
	LSD6	.14	-.15	.01
	LSD7	.18	.03	.69
	LSD8	.16	.06	.47
	LSD9	.11	-.03	.55
	LSD10	.09	-.00	.93
	LSD11	.07	-.01	.71
	LSD12	.25	.04	.46
	LSD13	.41	.17	.00
	LSD14	.29	.09	.14
	LSD15	.35	.17	.01
	LSD16	.35	.14	.03
	Multiple R	.5969	R Square	.3563
	Adjusted R Square	.2843	Standard Error	.7825
	F = 4.9495 (df=17, 152) p=.000			

POS= Possession Satisfaction

LSG= life as a whole

Domain Satisfaction:

LSD1=job situation; LSD2=family life; LSD3=financial situation; LSD4=health

LSD5=education; LSD6=friends and associates; LSD7=leisure life; LSD8=neighborhood

LSD9=community; LSD10=spiritual life; LSD11=taxes; LSD12=environment (air/water quality)

LSD13= political/economic freedom and independence; LSD14=housing situation; LSD15= cultural life;

LSD16=social status (1=terrible: 7=delighted)

Hypothesis 1 (H1)

H1 states that overall life satisfaction is influenced by affect or satisfaction in all major life domains including the material domain. Thus, life satisfaction is significantly influenced by material satisfaction above and beyond the influence of satisfaction from other life domains. This hypothesis was tested in two ways. First, we tested the model as a whole (materialism =>

possession satisfaction => overall life satisfaction). (see Table 1).

The correlation between life satisfaction (LSG) and satisfaction with possessions (POS) was .28 ($p < .01$) and the LISREL coefficient was .27 ($p < .05$) with R-square of .072. These results indicate that, as hypothesized, satisfaction with material domain positively affects overall life satisfaction.

The second way of testing the hypothesis was

by regressing life satisfaction scores against satisfaction with 16 different life domains including the material domain. The results are shown in Table 2.

The results indicate that material satisfaction positively affects overall life satisfaction (regression estimate = .17, $p < .05$). Other domains significantly effecting overall life satisfaction for college students were education (LSD5), friends and associates (LSD6), political/economic independence (LSD13), cultural life (LSD15), and social status (LSD16). The R-square for this equation was .356. This finding is consistent with previous studies and with H1 (Leelakulthanit, Day, and Walters 1991). It confirms that satisfaction with the material life domain and other life domains significantly affects overall life satisfaction.

Hypothesis 2 (H2)

H2 states that the influence of material satisfaction on life satisfaction is stronger for materialistic than for non materialistic individuals. This hypothesis was tested in two ways. First, we divided the sample into high and low materialism groups based on a median split. The median was 4.286. The results was 132 subjects were treated as materialistic (materialism score was greater than 4.286) and 151 subjects were treated as nonmaterialistic (materialism score was less than 4.286).

We regressed life satisfaction scores against satisfaction with the material domain with respect to each group. For the materialistic group, the .32 ($p < .01$) coefficient estimate between life satisfaction and satisfaction with possessions accounted for approximately 11 percent of the variance in life satisfaction. For the nonmaterialistic group, the .22 ($p < .05$) coefficient estimate between life satisfaction and satisfaction with possessions accounted for approximately 4.9 percent of the variance in life satisfaction (see Table 1). As hypothesized in H2, the spillover effects accounted for a greater percentage of the variance in life satisfaction for the high materialism group than for the low materialism group.

The second way of testing the hypothesis was by regressing life satisfaction scores of each group

(materialistic/nonmaterialistic) against satisfaction with 16 different life domains including the material domain. The results are shown in Table 3.

For the materialistic group, economic independence and material satisfaction significantly affected overall life satisfaction. For the nonmaterialistic group, education (LSD5), friends and associates (LSD6), leisure life (LSD7), and cultural life (LSD15) were significantly affected overall life satisfaction. It is interesting to note that material satisfaction (POS) was a significant indicator of life satisfaction for the materialistic group but had no effect on the life satisfaction of the nonmaterialistic group. Materialistic people have a strong emotional involvement in the material possession domain. As H2 suggests, the spillover from material domain to overall life satisfaction tends to be stronger for materialistic people than for nonmaterialistic people.

Hypothesis 3 (H3)

H3 stated that material satisfaction is negatively influenced by materialism. That is, those who are more materialistic are likely to express greater dissatisfaction with their possessions than those who are less materialistic. This hypothesis was tested in three ways. First, using LISREL, the following model was tested:

MATERIALISM -> SATISFACTION WITH POSSESSIONS -> LIFE SATISFACTION

The LISREL results produced coefficients indicative of a model with good-fit [$\chi^2 = 8.45$, $df = 1$, $p = .003$; RMR = .06; GFI = .98; AGFI = .89; NFI = .72; and CFI = .73]. The estimate of materialism on satisfaction with possessions was -.00 ($p > .10$), accounting for 7.2 percent of the variance in satisfaction with possessions (see Table 1). Though the relationship between materialism and possession satisfaction was not significant for the entire group (estimate = -.00, $p < .10$), there was a clear interaction when it was analyzed in the context of high vs. low materialism groups. The LISREL analysis showed a negative relationship (estimate = -.15, $p < .10$) for the high materialism group and a positive relationship (estimate = .21 $P < .05$) for nonmaterialistic group.

Table 3
Regression Analysis for High vs. Low Materialism Group
(Dependent Variable = Overall Life Satisfaction)

Sample	Independent Variable	Pearson Correlation	Regression Estimate	p-value
High Materialism (n=132)				
	POS	.35	.29	.04
	LSD1	.22	-.00	.94
	LSD2	-.03	-.05	.40
	LSD3	.05	-.04	.64
	LSD4	.11	.05	.56
	LSD5	.32	.00	.96
	LSD6	.11	.00	.99
	LSD7	-.03	-.07	.57
	LSD8	-.03	-.05	.74
	LSD9	.09	.07	.41
	LSD10	-.06	.03	.66
	LSD11	.02	-.06	.40
	LSD12	.13	-.04	.66
	LSD13	.51	.23	.01
	LSD14	.23	-.05	.66
	LSD15	.27	.12	.28
	LSD16	.36	.14	.16
	Multiple R	.6232	R Square	.3883
	Adjusted R Square	.2091	Standard Error	.7460
	F = 2.166 (df=17,58)	p=.0152		
Low Materialism (n=151)				
	POS	.22	.04	.73
	LSD1	.16	.00	.93
	LSD2	.20	.09	.28
	LSD3	-.03	-.11	.22
	LSD4	.21	.04	.65
	LSD5	.26	.21	.03
	LSD6	.15	-.17	.06
	LSD7	.42	.30	.05
	LSD8	.30	.12	.36
	LSD9	.16	-.16	.11
	LSD10	.29	.02	.80
	LSD11	.15	-.12	.21
	LSD12	.32	.10	.30
	LSD13	.33	.13	.15
	LSD14	.30	.12	.13
	LSD15	.39	.22	.03
	LSD16	.35	.08	.43
	Multiple R	.6890	R Square	.4747
	Adjusted R Square	.3453	Standard Error	.8113
	F = 3.6684 (df=17, 69)	p=.0001		

A second, more conservative test involves the same LISREL analysis; however, satisfaction with other life domains was also included as additional predictors of life satisfaction. These results are

shown in Tables 4 and 5.

This more conservative analysis produced a pattern similar to the first analysis. It showed that the effect of materialism on material satisfaction

Table 4
Lisrel Estimates with 16 Life Domains

Sample	Dependent Variable	Independent Variables	LISREL Estimate	T-value	R-Square
ENTIRE (n=297)					
	POS	IPOS	-.00	-.01	.00
	LSG	POS	.15*	3.10	.33
		LSD1	.06	1.13	
		LSD2	.01	.30	
		LSD3	-.10	-.16	
		LSD4	-.00	-.13	
		LSD5	.13*	2.08	
		LSD6	-.20*	-3.06	
		LSD7	.03	.53	
		LSD8	.07	1.0	
		LSD9	-.032	-.54	
		LSD10	.03	.55	
		LSD11	-.05	-.89	
		LSD12	.06	1.11	
		LSD13	.25*	4.04	
		LSD14	.12*	1.97	
		LSD15	.20*	3.44	
		LSD16	.16*	2.75	

($\chi^2 = 87.14$, $df=17$, $p=.00$, $GFI=.97$, $AGFI=.70$, $NFI=.95$, $CFI=.95$, $RMR=.04$)

* $p < .05$

was not significant for the entire group (estimate = .00, $p > .90$) and was significantly positive for the nonmaterialistic group (estimate = .24, $p < .05$). However, the effect of materialism on material satisfaction was negative (estimate = -.15, $p < .10$) for the materialistic group. This implies that those with high materialism tend to have high expectations which result in low satisfaction with the material life domain. Taken altogether, these results support H3 but only for high materialism subjects.

LIMITATIONS AND FUTURE DIRECTIONS

This study identified possession satisfaction as an antecedent of overall life satisfaction. That is, the satisfaction from the material possession life domain did spill over onto overall life satisfaction. This study has also found that the spillover effects are greater for those who are materialistic than for

those who are not. Finally, this study found that high materialism reduces satisfaction with the possession life domain for those who are highly materialistic.

This paper has certain limitations. First, we used a convenience sample of undergraduate students. It is, therefore, unclear whether the study's findings may be generalized to the entire population. A wider sample might, for instance respond to some items that were not relevant to students (e.g., satisfaction with house; satisfaction with investment and savings; job; taxes). Future studies may address this limitation by surveying subjects at other life stages. Second, this study focused on only one aspect of consumer life domain (possessions). The consumer life domain includes other related domains (e.g., acquisition, consumption, disposition etc.). Future work needs to be done to clarify the relationship between the other aspects of consumer life domain and overall

Table 5
Lisrel Estimates with 16 Life Domains:
High vs. Low Materialism Group

Sample	Dependent Variable	Independent Variables	LISREL Estimate	T-value	R-Square	
HIGH (n=132)	POS	IPOS	-.15	-1.66	.02	
		LSG	.26*	3.53	.36	
			LSD1	-.01	-.10	
			LSD2	-.10	-1.18	
			LSD3	-.07	-.65	
			LSD4	.07	.82	
			LSD5	.00	.05	
			LSD6	.00	.01	
			LSD7	-.00	-.80	
			LSD8	-.06	-.46	
			LSD9	.11	1.15	
			LSD10	.05	.61	
			LSD11	-.10	-1.17	
			LSD12	-.05	-.62	
			LSD13	.38*	3.63	
			LSD14	-.07	-.70	
			LSD15	.15	1.56	
		LSD16	.19*	2.00		
($\chi^2 = 76.69$, $df=17$, $p=.00$, $GFI=.95$, $AGFI=.45$, $NFI=0.92$, $CFI=.93$, $RMR=.04$)						
LOW (n=151)	POS	IPOS	.24*	2.85	.05	
		LSG	.03	.59	.47	
			LSD1	.01	.12	
			LSD2	.13	1.51	
			LSD3	-.14	-1.7	
			LSD4	.05	.63	
			LSD5	.25*	3.20	
			LSD6	-.23*	-2.63	
			LSD7	.28*	2.76	
			LSD8	.12	1.31	
			LSD9	-.19*	-2.27	
			LSD10	.03	.34	
			LSD11	-.17	-1.80	
			LSD12	.13	1.45	
			LSD13	.17*	2.01	
			LSD14	.17*	2.11	
			LSD15	.26*	3.06	
		LSD16	.09	1.1		
($\chi^2 = 61.02$, $df=17$, $p=0.00$, $GFI=.97$, $AGFI=.62$, $NFI=.94$, $CFI=.95$, $RMR=.04$)						

* $p < .05$

life satisfaction. Third, the role of materialism can be a function of culture (Leelakulthanit, Day and Walters 1991). Cross-cultural studies need to be done to identify what effect, if any, materialism may have when its absolute value is at a level higher or lower than that which is typical with this

sample of American college students. Finally, we acknowledged that while horizontal spillover may occur, we did not explicitly study this phenomenon. More research needs to be conducted on the impact of the material life domains on other life domains such as the

spiritual, health, or occupational domains.

These limitations notwithstanding, this study adds to our understanding of effects materialism has on the relationship between possession satisfaction and overall life satisfaction. The knowledge gleaned from this study can aid in psychographic segmentation, especially for materialistic consumers. The enhancement of QOL or life satisfaction is a major goal of QOL research. By identifying the conditions that affect the spillover from possession satisfaction to overall life satisfaction, this study can help policy makers develop programs that enhance overall life satisfaction. The policy implications of this study are as follows. Policy makers need to develop programs that enhance consumers' possession satisfaction. They might focus on those who are high in materialism because materialistic people tend to be more dissatisfied in the possession life domain than other people are. They also tend to have a high degree of spillover from the possession domain to overall life satisfaction.

REFERENCES

- Aiello Jr., A., J. A. Czepiel, and L. J. Rosenberg (1977), "Scaling the Heights of Consumer Satisfaction: An Evaluation of Alternative Measures," *Consumer Satisfaction, Dissatisfaction, and Complaint Behavior*, R. Day (ed.), Bloomington, IN: Indiana University School of Business, 43-50.
- Andrews, Frank M. and Stephen B. Withey (1976), *Social Indicators of Well-Being: America's Perception of Life Quality*, New York: Plenum Press.
- Belk, Russell W. (1983), "Worldly Possessions: Issues and Criticisms," *Advances in Consumer Research*, 10, Richard Bagozzi and Alice M. Tybout, (eds.), Ann Arbor, MI: Association for Consumer Research, 514-519.
- Belk, Russell W. (1984), "Three Scales to Measure Constructs Related to Materialism: Reliability, Validity and Relationships to Measures of Happiness," *Advances in Consumer Research*, 11, Thomas Kinnear, (ed.), Ann Arbor, MI: Association for Consumer Research, 291-297.
- Belk, Russell W. (1985), "Materialism: Trait Aspects of Living in the Material World," *Journal of Consumer Research*, 12, 265-280.
- Belk, Russell W. and Richard Pollay (1985), "Images of Ourselves: The Good Life in Twentieth Century Advertising," *Journal of Consumer Research*, 11(March), 887-897.
- Bloch, Peter H. and Marsha L. Richins (1983), "A Theoretical Model for the Study of Product Importance Perceptions," *Journal of Marketing*, Summer, 69-81.
- Brewer, W. and G. V. Nakamura (1984), "The Nature and Functions of Schemas," *Handbook of Social Cognition*, R. S. Wyer Jr. and T. K. Srull, (eds.), Hillsdale, NJ: Erlbaum, 119-159.
- Brief, Arthur P., Ann H. Butcher, Jennifer M. George and Karen E. Link (1993), "Integrating Bottom-Up and Top-Down Theories of Subjective Well-Being: The Case for Health," *Journal of Personality and Social Psychology*, 64 (4), 646-653.
- Burke, P. L. and J. C. Tully (1977), "The Measurement of Role Identity," *Social Forces*, 55, 881-897.
- Campbell, A., P. E. Converse and W. L. Rodgers (1976), *The Quality of American Life: Perceptions, Evaluations, and Satisfaction*, New York: Russell Sage Foundation.
- Celsi, Richard L. and Jerry C. Olson (1988), "The Role of Involvement in Attention and Comprehension Processes," *Journal of Consumer Research*, September, 210-214.
- Cole, Dennis, Newell D. Wright, M Joseph Sirgy, Rustan Kosenko, Don Rahtz and H. Lee Meadow (1992), "Testing the Reliability and Validity of Two Measures of Materialism," *Developments in Marketing Science*, 15, Victoria L. Crittenden, (ed.), Chestnut Hill, MA: Academy of Marketing Science, 383-387.
- Dawson, Scott and Gary Bamossy (1990), "Isolating the Effect of Non-Economic Factors on the Development of Consumer Culture: A Comparison of Materialism in the Netherlands and the United States," *Advances in Consumer Research*, Marvin Goldberg, Gerald Gorn, and Richard W. Pollay, (eds.), Provo, UT: Association for Consumer Research, 182-185.
- Dawson, Scott and Gary Bamossy (1991). "If 'We Are What We Have,' What Are We When We Don't Have To Have Possessions? A Handbook of Ownership and Property, Special Issue, *Journal of Social Behavior and Personality*, 6(6), 363-384.
- Day, Ralph L. (1978), "Beyond Social Indicators: Quality of Life at the Individual Level," *Marketing and the Quality of Life*, Fred D. Reynolds and Hiram C. Barksdale, (eds.), Chicago, IL: American Marketing Association, 11-18.
- Day, Ralph L. (1987), "Relationship between Life Satisfaction and Consumer Satisfaction," *Marketing and Quality-of-Life Interface*, A. C. Samli, (ed.), Westport, CT: Greenwood Press, 289-311.
- Diener, Edward (1984), "Subjective Well-Being," *Psychological Bulletin*, 75(3), 542-575.
- Douglas, M. and B. Isherwood (1979), *The World of Goods*, New York: Basic.
- Douthitt, Robin A., Mauric MacDonald and Randolph Mullis (1992), "The Relationship Between Measures of Subjective and Economic Well-Being: A New Look," *Social Indicators Research*, 26, 407-422.
- Easterlin, Richard A. and Eileen M. Crimmins (1988),

- "Recent Social Trends: Changes in Personal Aspirations of American Youth," *Sociology and Social Research*, 72 (July), 217-223.
- Easterlin, Richard A. and Eileen M. Crimmins (1991), "Private Materialism, Personal Self-Fulfillment, Family Life, and Public Interest: The Nature, Effects, and Causes of Recent Changes in the Values of American Youth," *Public Opinion Quarterly*, 55 (Winter), 499-533.
- Fishbein, Martin and Icek Ajzen (1975), *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*, Reading, MA: Addison-Wesley.
- Garza, R. T. and L. G. Herringer (1987), "Social Identity: A Multidimensional Approach," *Journal of Social Psychology*, 127, 299-308.
- Griffin, N. L. Chassin, and R. Young (1981), "Measurement of Global Self-Concept Versus Multiple Role-Specific Self-Concepts in Adolescents," *Adolescence*, 16, 49-56.
- Hoelter, J. W. (1985), "The Structure of Self-Conception: Conceptualization and Measurement," *Journal of Personality and Social Psychology*, 49, 1392-1407.
- Houston, M. J. and M. L. Rothchild (1977), "A Paradigm for Research on Consumer Involvement," Working Paper No. 11-77-46, University of Wisconsin-Madison.
- Iris, B. and G. V. Barrett (1972), "Some Relations Between Job and Life Satisfaction/Job Importance," *Journal of Applied Psychology*, 49, 112-126.
- Kleine, R. E., III and Jerome B. Kernan (1991), "Contextual Influences on the Meanings Ascribed to Ordinary Consumption Objects," *Journal of Consumer Research*, 18, 311-324.
- Kleine, R. E. III, S. S. Kleine and Jerome B. Kernan (1992), "Mundane Everyday Consumption and the Self: A Conceptual Orientation and Prospects for Consumer Research," In *Advances in Consumer Research*, 19, J. F. Sherry and B. Sternthal, (eds.), Provo, UT: Association for Consumer Research, 411-415.
- Kleine, R. E. III, S. S. Kleine and Jerome B. Kernan (1993), "Mundane Consumption and the Self: A Social-Identity Perspective," *Journal of Consumer Psychology*, 2(3), 209-235.
- Larsen, R. J., R. A. Emmons, and Ed Diener (1983), *Validity and Meaning of Measures of Subjective Well Being*, Paper presented at the Midwestern Psychological Association Convention, Chicago.
- Leelakulthanit, Orose, Ralph Day and Rockney Walters (1991), "Investigating the Relationship between Marketing and Overall Satisfaction with Life in a Developing Country," *Journal of Macromarketing*, (Spring), 3-23.
- Markus, H. (1977), "Self-Schemata and the Processing of Information about the Self," *Journal of Personality and Social Psychology*, 35, 63-78.
- McCall, G. J. and J. L. Simmons (1978), *Identities and Interactions: An Examination of Human Associations in Everyday Life*, New York: Free Press.
- McCracken, G. (1988), *Culture and Consumption*. Bloomington: Indiana University Press.
- Morgan, A. J. (1993), "The Evolving Self in Consumer Behavior: Exploring Possible Selves," *Advances in Consumer Research*, 20, L. McAlister and Rothchild, (eds.), Provo, UT: Association for Consumer Research, 429-432.
- Meadow, H. Lee (1988), "The Satisfaction Attitude Hierarchy: Does Marketing Contribute?" *Proceedings of the 1988 American Marketing Association Winter Educators' Conference*, S. Shapiro and others, (eds.), Chicago, IL: American Marketing Association, 482-483.
- Neugarten, B. L., R. J. Havighurst and S. S. Tobin (1961), "The Measurement of Life Satisfaction," *Journal of Gerontology*, 16, 134-143.
- Richins, Marsha L. (1987), "Media, Materialism, and Human Happiness," *Advances in Consumer Research*, 14, Melanie Wallendorf and Paul Anderson, (eds.), Provo, UT: Association for Consumer Research, 352-356.
- Richins, Marsha L. and Scott Dawson (1992), "A Consumer Values Orientation for Materialism and Its Measurement: Scale Development and Validation," *Journal of Consumer Research*, 19(December), 303-316.
- Rosenberg, M. (1978), *Conceiving the Self*, New York: Basic Books.
- Rudmin, Floyd and Marsha Richins (1992), *Meaning, Measure, and Morality of Materialism*, Provo, UT: Association for Consumer Research.
- Seeman, M. (1967), "On the Personal Consequences of Alienation of Work," *American Sociological Review*, 32, 273-285.
- Scott, W. A. and J. Stumpf (1984), "Personal Satisfaction and Role-Performance: Subjective and Social Aspects of Adaptation," *Journal of Personality and Social Psychology*, 47, 812-827.
- Sirgy, M. Joseph, H. Lee Meadow and A. C. Samli (1995), "Past, Present, Future: An Overview of Quality-of-Life Research in Marketing," *New Dimensions in Marketing/Quality-of-Life Research*, M. Joseph Sirgy and A. C. Samli, (eds.), Westport, Connecticut: Quorum Books, 335-364.
- Sirgy, M. Joseph, David E. Hansen, and James E. Littlefield (1994), "Does Hospital Satisfaction Affect Life Satisfaction?" *Journal of Macromarketing*, 14(Fall), 36-46.
- Stock, W., M. Okun, S. Stock, M. Haring and R. A. Winter (1982), *Reporting Reliability: A Case Study of Life Satisfaction Research*, Paper presented at the 90th Annual Convention of the American Psychological Association, Washington, DC.
- Stryker, S. (1968), "Identity Salience and Role Performance: The Relevance of Symbolic Interaction

Theory for Family Research," *Journal of Marriage and the Family*, 30, 558-564.

Veenhoven, R. (1991), "Is Happiness Relative?" *Social Indicators Research*, 24, 1-34.

Wright, Newell and Val Larsen (1993), "Materialism and Life Satisfaction: A Meta-Analysis," *Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior*, 6 158-165.

Send correspondence regarding this article to:

Newell Wright
Department of Marketing and HTM
James Madison University
Harrisonburg, VA 22807 USA

"AFTER I HAD MADE THE DECISION, I...:" TOWARD A SCALE TO MEASURE COGNITIVE DISSONANCE

Douglas R. Hausknecht, University of Akron
Jillian C. Sweeney, University of Western Australia
Geoffrey N. Soutar, Edith Cowan University
Lester W. Johnson, Monash University

ABSTRACT

The objective of the present research study is to begin the development of a scale to measure cognitive dissonance that arises from free choice in consumer purchase decisions. This represents an advance in the measurement of cognitive dissonance. Cognitive, emotional and behavioral aspects of the dissonance construct are isolated and described, as are relationships with other post purchase variables. In particular, it is noted that the reduction of cognitive dissonance is a necessary condition for the occurrence of satisfaction. Multiple items, developed from focus groups, were evaluated by a series of independent judges (consumer behavior researchers). A total of 81 measurement items are offered for further refinement.

INTRODUCTION

"One hopes that the construction, validation and dissemination of comprehensive dissonance scales will be forthcoming." (Oliver 1997, p. 261) With this plea, Oliver concluded his chapter length review of dissonance research as part of a landmark treatise on consumer satisfaction. Since Leon Festinger coined the term "cognitive dissonance" in 1957, the concept has been interpreted, debated, and re-interpreted with some frequency and ferocity. Many of the early critical issues, however, remain at the center of disagreements. Underlying these continued debates are attempts to fit the same concept into a set of evolving theories and paradigms.

Various conceptualizations have been used in attempts to identify the relationships between cognitive dissonance and consumer satisfaction/dissatisfaction [CS/D]. Some authors have discussed how dissonance generally fits theories of consumer behavior (e.g., Cummings and Venkatesan 1976; Schewe 1973), while others have made explicit distinctions among the constructs and their relationships (e.g.,

Montgomery and Barnes 1993; Oliver 1997). To date, however, no one has settled convincingly the conflicts that have been raised.

Whereas the satisfaction construct has been widely discussed, and measures and models developed around it (e.g., Churchill and Surprenant 1982; Johnson and Fornell 1991; Oliver 1980), fewer measurement studies have examined the concept of dissonance. Some early dissonance studies have been criticized as tapping related constructs, such as anxiety, rather than dissonance itself (Cummings and Venkatesan 1976). Other earlier studies did not measure dissonance but, rather, inferred the occurrence of dissonance from evidence of dissonance reducing behaviors. For example, Engel (1963) assessed the attention paid to advertising of a brand after purchasing that brand (Engel 1963), while Loscuito and Perloff (1967) measured attitude changes towards selected and non-selected products. Both are examples of how people strive to achieve congruence among attitudes, knowledge and behavior (Schewe 1973).

Despite four decades of discussion, an agreed upon and measurable construct continues to elude social scientists. It has been suggested recently that, "dissonance, at least as presently measured, may not have discriminant validity when compared to other post purchase constructs" (Sweeney, Soutar and Johnson 1996, p. 138). Oliver (1997) discusses the relationships among these constructs with an acknowledgment of the need for improved measures consistent with theoretical models.

THE DISSONANCE CONCEPT

Festinger's early explanation of the dissonance construct does not identify clearly whether "dissonance" is cognitive or emotional. The cognitive view is supported by his definition that "the obverse of one element follows from the other" (Festinger 1957, p. 261). Yet, he seems also to have intended an emotional conceptualization, suggesting that "for some

people, dissonance is an extremely painful and intolerable thing" (Festinger 1957, p. 266). In a recent social psychology treatment, Gilovich, Medvec and Chen (1995), appear to treat dissonance in a traditional cognitive sense. Conversely, Harmon-Jones et al. (1996) seem to be more concerned with the emotions aroused by the "aversive consequences" of an individual's action. In fact, they subscribe to the theory that it is these consequences that are necessary and sufficient to create the emotion, rather than mere cognitions. However, they were working within the "induced-compliance" paradigm and their conclusions may not have direct relevance to the "free-choice" paradigm that is most often of interest to consumer researchers.

Oliver (1997) revived the free-choice version of dissonance, which he characterized as having lain dormant for some time. In Oliver's model of satisfaction, the dissonance concept is stretched across two-thirds of the satisfaction process. Originating in a pre-purchase phase, the construct is labeled "apprehension." These same cognitions and feelings mutate into true dissonance after the decision is made. With use and experience, dissonance dissipates and yields to dis/satisfaction (as can be seen in Oliver's figure 1-3, 1997, p. 22).

While Oliver (1997) argues that dissonance occurs at various stages of the consumption process, it is generally recognized as a post-decisional, but pre-use phenomenon (e.g. Festinger, 1957; Insko and Schopler, 1972). Indeed, Oliver (1997, p. 24), in a subsequent section, views dissonance as resulting "from a personal decision or action." He termed this narrower window the "Gamma" phase (Oliver 1997, p. 242). The concept of dissonance addressed in the present research best fits this period that immediately follows the purchase decision but precedes use or experience with the result of the purchase decision.

This relationship is made explicit in Figure 1, in which the horizontal axis represents changes over a purchase and consumption process, but does not presume causality. Dissonance constructs arise only after the decision is made and in response to a number of factors.

It is important to note that the presentation of satisfaction models and dissonance models is not

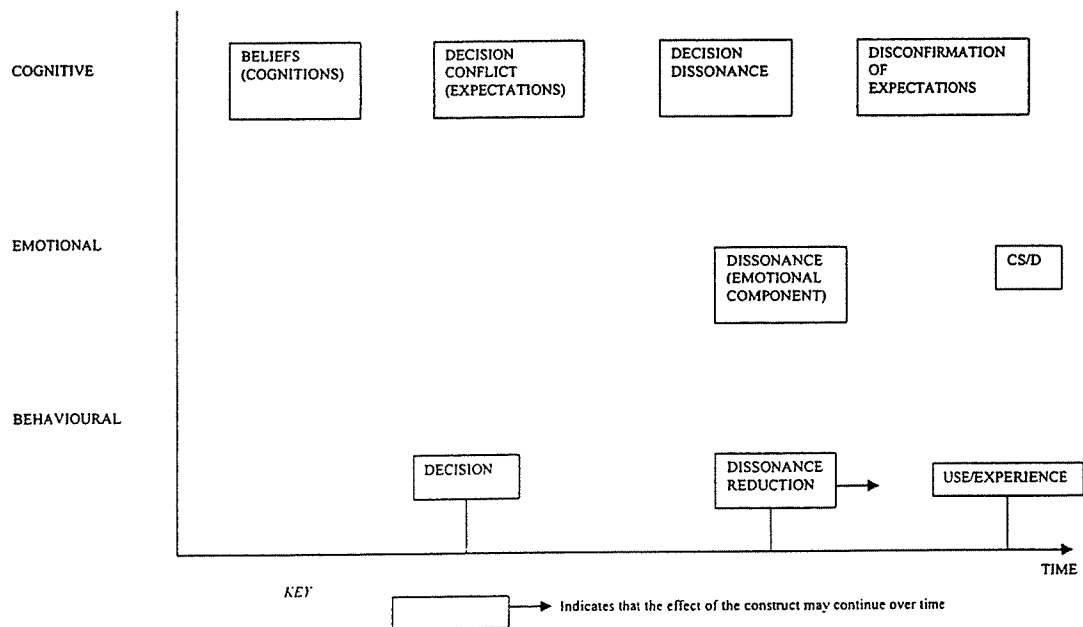
meant to pre-suppose that every purchase results in the arousal of either or both of these processes. For example, it has been argued that satisfaction and/or dissatisfaction may not arise in low involvement situations (Hausknecht 1988; Oliver 1997). Using Oliver's (1997, p. 13) concept of satisfaction as a fulfillment response, it is apparent that either the cognitive or emotional components may not be aroused in given situations.

Bell (1967) suggested long ago that some individuals simply may not experience dissonance. The literature has established well the *necessary* conditions for dissonance arousal (Oliver 1997); although there is some confusion caused by mixing paradigms (i.e., forced compliance versus free choice). A physiological state, arousal, has been suggested as another necessary condition (Elliot and Devine 1994), but no one has demonstrated *sufficient* conditions to force the process. The forced compliance paradigm has been used most often in dissonance studies to ensure having something to measure, but has been assailed as not likely, or even rare, in consumption situations (Cummings and Venkatesan, 1976; Oliver 1997).

The previously mentioned controversy as to the treatment of dissonance as cognitive or emotional is similar to that which exists in the attitude literature between proponents of the tripartite attitude model and those who favor the attitude-as-affect version. To clarify the nature of the constructs, the present model introduces distinctions over time and separates cognitive from emotional concepts. The first of these, decision conflict, is not usually presented as part of the consumer decision process. Davidson and Kiesler (1964) cited Festinger's distinction between decision conflict, a pre-decision concept, and dissonance, a post-decision concept. The same authors also present a contrary view, "Janis (1959) on the other hand, thinking in terms of 'conflict resolution', implied that there is little or no distinction between pre- and post-decision behavior and that systematic re-evaluation occurs both before and after the decision" (Davidson and Kiesler 1964, p. 10). Even if the two concepts are similar structurally, they are divided in time by the act (behavior) of having made a decision. In any case, they describe decision conflict as a cognitive imbalance that is resolved by the decision.

Whatever the relationship of dissonance with

Figure 1
Temporal Relations Among Dissonance Concepts



pre-decision variables may be, there is consensus that beliefs that are inconsistent with a decision may persist after a purchase. For clarity of exposition, Figure 1 labels this 'decision dissonance'. If decision conflict exists up to the moment of decision, it is decision dissonance that follows immediately. Insko and Schopler (1972, p. 109) reasoned similarly, suggesting that "dissonance is thus postdecisional conflict." Further, they suggested that there may be a "spreading" of the evaluation of the decision alternatives as part of the conflict or the dissonance. Oshikawa (1972, p. 65) agreed with the temporal positioning of dissonance but labeled it as only, "an intervening variable; ... after a decision is made and before attempts at dissonance reduction."

More substantially, this period in a decision process has been characterized as representing a change in confidence (Knox and Inkster 1968), a motivation to solve a puzzle regarding one's own behavior (Greenwald and Ronis 1978) or a wonder about the wisdom of the decision (Lowe and

Steiner 1968). All of these concepts appear to be devoid of feeling, evaluation, or emotion. Thus, decision dissonance appears to be the same kind of mental entity as cognitions and decision conflict.

Cognitive dissonance is most commonly defined as psychological discomfort (Carlsmith and Aronson 1963; Elliot and Devine 1994; Festinger 1957), a psychologically uncomfortable state (Menasco and Hawkins 1978), linked with anxiety and uncertainty or doubt (Montgomery and Barnes 1993; Mowen 1995; Oshikawa 1972) or synonymous with the regret or remorse reported in salespeople's anecdotes (Insko and Schopler 1972; Mowen 1995). Thus, the forty-year history of the literature appears to have created a theoretical oxymoron in which an essentially emotional construct bears the burden of "cognitive" in its name. Nevertheless, "to date, there have been no systematic attempts to directly empirically validate the psychological discomfort component of dissonance" (Elliot and Devine 1994, p. 383). The emotional aspect of dissonance, representing the psychological discomfort

described above, is termed "dissonance - emotional component" in Figure 1. Subsequent to the arousal of cognitive dissonance, dissonance reduction behaviors may arise and these have often been used as surrogate measures for dissonance in prior work. Finally, satisfaction is assessed post-purchase and post-use, when performance is compared to expectations. Satisfaction has been described as an affective (emotion) state or feeling reaction, in which the consumer's needs, desires and expectations during the course of the consumption experience have been met or exceeded (Oliver, 1989). Dissonance, should it be induced, is critical to the formation of satisfaction, in that dissonance reduction is a necessary condition for satisfaction to occur (Oliver, 1997).

Based on the foregoing definitions and distinctions, this paper seeks to identify measures that can be used to measure cognitive dissonance discreetly from its related constructs. In the next section, prior measures are reviewed briefly and the first steps are taken to identify candidate items for a final measure.

MEASURES

As is so often true in the consumer behavior literature, inconsistencies in construct definition and measurement go hand in hand. A variety of techniques have been used to measure dissonance and these are not always consistent with the definition asserted by the researcher(s). One way to view these measures is along the same cognitive versus emotional versus behavioral dimensions used previously for CS/D measures (Hausknecht 1990). These dimensions were used for the purposes of the present study, which considers dissonance as occurring post-purchase, to disentangle the concept as follows:

Cognitive Construct - a person's recognition that beliefs which are inconsistent with a decision exist after the purchase has been made

Emotional Construct - a person's psychological discomfort subsequent to the purchase decision

Behavioral Construct - a person's actions that

serve to confirm a decision or reduce tension

These classifications approximate the concepts of decision dissonance, dissonance (emotional component), and dissonance reduction described above. The first two classifications also correspond to the antecedents of dissonance and psychological aspects of dissonance itself described by Oliver (1997).

Some examples of these measures from the existing literature are:

Cognitive measures

a. To what extent do you wonder whether or not you made the right decision?

Would most people expect to get the same kind of deal you got?
(Bell 1967)

b. I feel that I will be happy with the purchase I have just made.

I'm sure that I'll be pleased with the way this product performs.

I'm confident that I've made the "right" choice when I purchased this product.
(Montgomery and Barnes 1993)

c. I should have spent more time shopping.

It was difficult to decide which brand to buy.

(Menasco and Hawkins 1978; Sweeney, Soutar and Johnson 1996)

d. Differences in the valuation (cognition) of chosen versus unchosen alternatives.

(Gilovich, Medvec and Chen 1995)

Emotional measures include:

a. I am comfortable/uneasy with the purchase decision I've just made.

(Montgomery and Barnes 1993)

b. Measures of anxiety

(Bell 1967; Hunt 1970)

c. How do you feel right now?

(Uncomfortable, uneasy, bothered)

(Elliot and Devine 1994)

d. Similar to c above, but with the addition of anxious, tense, apprehensive. (Oliver 1997)

Behavioral measures include:

a. All of the attempts to measure or record behaviors thought to "reduce dissonance."

b. Behavioral intent scales, e.g.,
I'll probably talk to my friends or family to ask them if they think I've made a wise choice with my purchase.
I would probably pick up a copy of Consumer Reports to make sure the product or brand I just bought received high ratings.
(Montgomery and Barnes 1993)

Most researchers admit that the behavioral measures are, at best, indirect measures of dissonance. Under its earlier conceptualization, many were concerned about being able to measure the construct directly (e.g., Menasco and Hawkins 1978; Oshikawa 1972). In the mid 1970's, Cummings and Venkatesan (1976, p. 304) argued for a "theoretically relevant and unambiguous measure of magnitude of dissonance." By necessity, this should be as direct a measure as possible. The effort by Elliot and Devine (1994) is instructive, but not directly on point. Their experiments were grounded strictly in the forced compliance paradigm (students writing counter-attitudinal essays) and the conclusions reached are of dubious relevance to consumption situations. A scale of cognitive dissonance within the free-choice paradigm was developed by Montgomery and Barnes (1993, p. 206), who defined the domain of cognitive dissonance as "those feelings attitudes and emotions that consumers have or display when they experience dissonance and the situations and conditions in which dissonance has occurred." Thus, by definition, they included emotions as outward evidence of dissonance (e.g. "dissonant consumers often display anxiety", or "dissonant consumers may experience low levels of satisfaction"). This is not necessarily dissonance

itself. In addition, they included "support" in their measure, stating that "dissonant consumers need reassurance that a wise purchasing decision has been made" (Montgomery and Barnes 1993, p. 206-7). The latter is a method of dissonance reduction, in contrast to dissonance itself. Further, items used in their 16 item scale were derived from the marketing literature, rather than from consumers relating dissonance experiences. While the Montgomery and Barnes (1993) scale is a meritorious attempt to clarify this elusive construct, the present research generates measures of dissonance from consumers' thoughts and feelings following a difficult decision. Emotional and cognitive components of dissonance only will be pursued in the present study due to the inappropriateness of behavioral measures, as already discussed.

METHOD FOR DEVELOPING A MEASURE OF COGNITIVE DISSONANCE

For the present study, we turned to consumers to generate descriptions of thoughts and feelings at the time of a purchase decision. As Churchill (1979) has noted, this is best achieved through exploratory designs. A combination of such designs was used to elicit the items, which are to be purified in later stages.

Four focus groups were held with consumers in a large metropolitan area in Western Australia. Two sessions were conducted using typical focus group procedures while two used an emerging computerized technique that has been labeled Group Support Systems or GSS (for a full description of this approach see, Soutar, Whiteley and Callan 1996). A total of 34 consumers participated, none in more than one session. Participants were recruited by a professional marketing research firm and were paid a modest honorarium.

The questions used and scenarios given were similar across all groups. Each participant was asked at the start of each session to identify a recent important purchase decision that involved a difficult choice (which may have been among two or more close alternatives). Products selected by respondents varied from expensive shoes to cars and blocks of land. Once each decision was identified, some discussion of the incidents ensued

MODELING THE IMPACT OF RETAIL STOCKOUTS: IMPLICATIONS FOR CUSTOMER SATISFACTION AND RETENTION

Carlo D. Smith, The University of Tennessee

ABSTRACT

The cost of retail out of stock conditions extends beyond the immediate loss in sales revenue. The impact of stockout conditions is reflected in customer satisfaction, loyalty, and resulting shopping behaviors. A conceptual model is presented to describe customer shopping behaviors prompted by stockout conditions and how such conditions may affect customer satisfaction and loyalty toward the service provider.

INTRODUCTION

A 1996 Andersen Consulting study "The Retail Problem of Out-of-Stock Merchandise" offered dramatic evidence of the continuing significance of inventory stockout problems discussed by Schary and Christopher in 1979. The Andersen study, conducted in the retail grocery industry, presented the following summary results:

8.2% of items in the categories studied were out of stock each day, 11% on Sundays,

15% of all advertised items were out of stock on a daily basis during promotions,

66% of lost sales due to out of stock positions involved items which contributed to the fastest moving 25% of items,

34% of the time that products were out of stock, consumers postponed their purchases or looked elsewhere,

46% of the sales dollars associated with out of stock items were lost,

\$7 to \$12 billion of sales was "up for grabs" each year because merchandise was out of stock on store shelves.

While it is apparent that the immediate impact of a retail stockout may include the loss of revenue because customers cannot make their intended

purchase and choose not to buy an alternative, the longer term implications for corporate performance include the potential reduction in customer patronage, loyalty, and associated future revenues, as well the threat that negative word of mouth may deter other current and potential customers from patronizing the retailer (Schary and Christopher 1979).

Since the Schary and Christopher article, sophisticated point-of-sale and information management systems have provided retailers with more effective mechanisms to estimate the immediate impact of stock related lost sales. The impact of out-of-stock positions on customer loyalty, retention, and word of mouth behavior, however, is more difficult to assess.

Little research has been conducted to evaluate the impact of retail stockouts on customer behavior (Emmelhainz, Stock, and Emmelhainz 1991). The topic is of concern to researchers and practitioners in logistics, retail management, and marketing. Each of these areas has focused on different issues related to stock out conditions.

Logistics literature has addressed planning and control issues important in determining statistically based stock objectives necessary to meet customer service goals (Ballou 1992), customer decisions and behaviors resulting from out of stock conditions (Walter and LaLonde 1975, Emmelhainz, Emmelhainz, and Stock 1991), and the assessment of the economic cost of lost sales due to stockout conditions (Walter and LaLonde 1975).

Retail literature has also investigated customer decision processes associated with stockout conditions (Emmelhainz, Stock, and Emmelhainz 1991, Schary and Christopher 1979) and incorporated stock selection into the assessment of retail customer service (Westbrook 1981). Yet with few exceptions (Schary and Christopher 1979), the implications of retail stockout conditions have been considered part of an aggregate retail service construct.

Marketing literature in consumer satisfaction and service quality has not addressed inventory stock availability unless it has been interpreted as a tangible attribute of service quality (Zeithaml,

Berry, and Parasuraman 1993). The primary venue of satisfaction literature has involved the evaluation of post purchase product satisfaction (Oliver 1989, Churchill and Surprenant 1982, Woodruff, Cadotte, and Jenkins 1983) and more recently, the evaluation of dis/satisfaction with service offerings (Dabholkar 1995a). Service quality literature, on the other hand, has primarily been concerned with the identification and measurement of attributes of the service quality construct and, more recently, the assessment of the impact of service improvements on corporate performance (Zeithaml, Berry, and Parasuraman 1996).

While some companies have been rewarded for efforts to improve service quality, others have been reluctant to invest in such improvement efforts without evidence of a profitable outcome (Zeithaml, Berry, and Parasuraman 1996). This concern highlights the importance practitioners place on identifying a more effective means for assessing consequences associated with variations in customer service.

The purpose of this paper is to offer insights into the mechanisms by which a change in retail stock conditions, more specifically a stockout position, may affect the customer's immediate evaluation of dis/satisfaction with the retail service provider. The framework also indicates the potential longer term implications for the development of customer attitudes toward retailer loyalty and their resulting behaviors involving future patronage. This discussion is prompted by the following research questions:

How do retail stockout conditions impact short-term customer dis/satisfaction?

How do repeated experiences associated with retail stockout conditions affect longer term customer loyalty and retention?

To address these questions, a conceptual model of customer reaction to stock availability is presented. The theoretical base for this model combines research from dis/satisfaction and service quality literature along with that covering logistics and retail management. The model is used to propose changes in attitudes, intentions, and behaviors resulting from retail inventory stockout

conditions. The dynamics of the model presented considers potential short-term as well as long-term changes in customer shopping behavior, loyalty, and patronage resulting from the evaluation of service at the immediate transaction level.

In this paper, we propose to extend research in two ways. First, we consider the application of objective measures in the evaluation of customer satisfaction and service quality. While numerous measures have been developed to assess consumer perceptions of satisfaction and service quality, there have been few efforts to identify and develop more objective measures such as stock availability. Previous research by Rust and Zahorik (1993) has quantified the financial impact of more objective customer satisfaction elements in the banking industry. Retail stock availability is proposed to offer one such attribute in the retail merchandise industry which may be monitored and which provides management with a means to effectively influence customer service in retail establishments.

Second, we address a secondary assessment of dis/satisfaction toward the retail outlet as a result of a stockout condition. Such an assessment occurs when a stockout position leads the customer to purchase a substitute product, (resulting in an assessment of dis/satisfaction with the product), or visit an alternate retail outlet, (resulting in an assessment of dis/satisfaction with the alternative outlet), and its influence on customer attitude toward the original retail outlet. This perspective on customer dis/satisfaction incorporates adaptation level theory (Oliver 1980) which views satisfaction as a source of influence regarding attitude change and purchase intention.

The discussion is presented in three sections. First, literature addressing the role of stock availability in retail customer service is presented. Second, a model of customer reaction to stock availability is introduced and explained. The model is presented in four components including customer satisfaction with retail outlet service quality, customer reaction to retail stockouts, the impact of alternate product selection on service expectations and attitude toward the primary retailer, and alternate service provider impact on service expectations and attitude toward the primary retailer. Finally, implications for research and practice are offered.

CUSTOMER REACTIONS TO STOCK AVAILABILITY

Much of the extant literature regarding customer reactions to out of stock conditions focuses on the immediate behavioral responses to out-of-stock positions.

In a 1991 study, Emmelhainz, Stock, and Emmelhainz manipulated five items to be out of stock at a discount grocery store. They identified three factors determining item selection including product attributes, product purchase frequency, and the availability of size, variety, and brand substitutes. The three most common actions taken when a preferred item was not available included substituting a different brand and keeping size and variety constant, substituting a different variety and keeping size and brand constant, or not making a substitution and planning to go to a different store. In addition they identified three factors which appeared to affect the decision whether or not to substitute an item: perceived product risk, intended product usage, and urgency of need.

In a survey of patrons of conventional department stores, Westbrook (1981) identified eight factors associated with store characteristics which may impact retail customer satisfaction including salesperson, store environment, store merchandising practices, service orientation, retailer related experiences, store clientele, value of offering, and advertised sales. Within this framework, stock availability would be considered an attribute of the third factor associated with merchandise selection. While this study focused on measurements related to the experience of being in the store and consuming the products and services obtained from the retailer, the survey questions focused on items purchased. This rules out any explicit measure of the impact of stockout conditions on customer satisfaction.

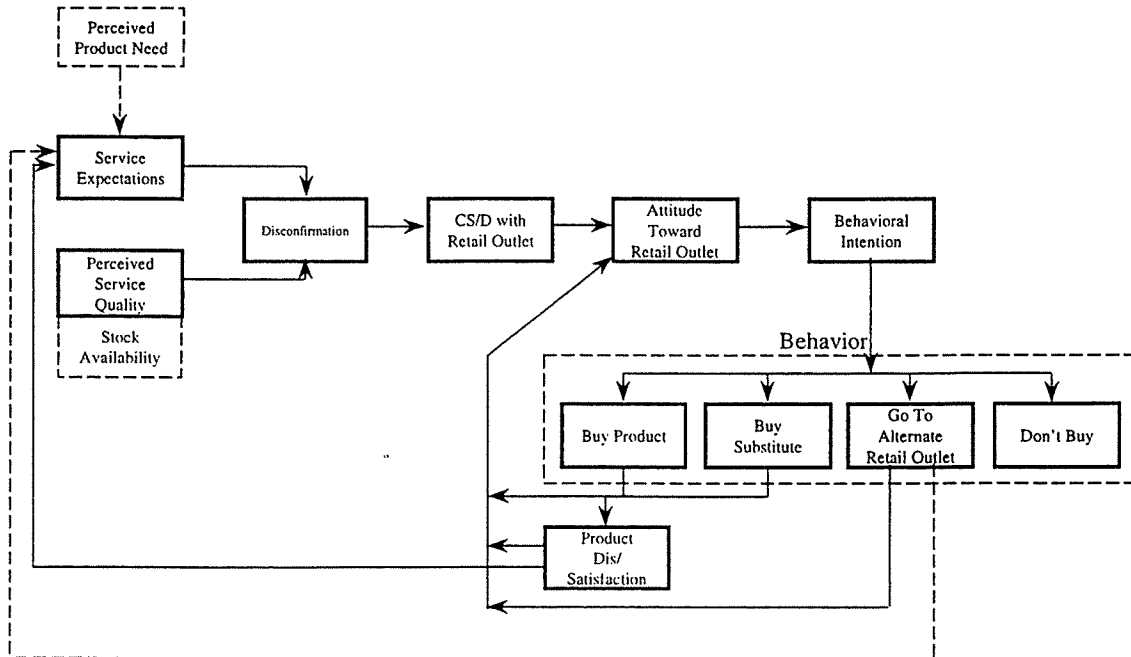
Schary and Christopher (1979) presented a detailed process model of customer stockout behavior and reported on the results of a survey conducted at two units of a British supermarket chain. Their model utilized some aspects of disconfirmation theory suggesting that consumer behavior presumes expectation based on previous performance or other sources of objective knowledge. Incorporating four component

modules associated with decisions in response to stockout conditions: store and product decisions, consumer behavior, response to out of stock situations, and retail merchandising strategies, only stockout behavior and merchandise strategy are suggested as offering an opportunity for measurement. Schary and Christopher noted a potential shift in customer attitude toward the retailer in addition to the behaviors displayed in response to the stockout conditions.

Walter and LaLonde (1975) and Walter and Grabner (1975) quantified the economic cost from retail out-of-stock conditions. They introduced and tested a stockout model which suggested six alternative customer behaviors in response to a stockout condition including purchasing a substitute at a higher price, purchasing a substitute at the same price, purchasing a substitute at a lower price, purchasing the same brand in a different size, returning to the store at a later date, or visiting an alternative store. By measuring the frequency of each behavior in addition to product costs, prices, and average sales per customer, the model established an expected value for an out-of-stock condition. They extended the model to consider behaviors associated with customers experiencing repeated stockout conditions after two visits to the retail outlet. These extensions included the possibility of placing a special order for the item or asking for the original items desired. The repeated stockout model presented the option of switching stores before determining a product substitution strategy. The "basic stockout model" as well as the "model for repeated stockouts" were tested in a survey of customers patronizing liquor stores in the state of Ohio. The Ohio state government controls and operates all retail liquor outlets in the state. This presented a unique opportunity to assess the implications of stockouts because of the limited pricing and promotion activities. Results indicated an expected revenue loss equal to 24 per cent of the retail price of the average product.

Walter and LaLonde (1975) and Walter and Grabner (1975) presented an effective method to assess the cost of lost sales associated with product stockout conditions. A potential limitation recognized by the authors involved the generalizability of the model to more competitive situations incorporating alternative retail sources

Figure 1
A Model of Customer Reactions to Stock Availability



and pricing factors. In addition, both stockout models focused on the immediate behavior of individuals patronizing the store and did not consider the impact on customer attitudes and future shopping behaviors.

MODELING THE CUSTOMER REACTION TO STOCK AVAILABILITY

Figure 1 presents a conceptual model of the retail shopping experience of an individual. More specifically, the model illustrates the temporal changes in attitude and behavioral intentions related to an evaluation of dis/satisfaction with retail service. The model also identifies alternative behaviors which may be exhibited by customers based on stock availability.

To summarize the process depicted, the model begins with a customer's perceived need for a product, leading to an assessment of service from the selected retail outlet based on availability of the item. If the product is available, service

expectations are confirmed, attitude toward the retail outlet is likely to remain stable or become more favorable, the customer purchases the desired product, and the focus of dis/satisfaction then is directed toward an assessment of the product at the time of consumption (as depicted in the Product Dis/Satisfaction construct). In cases where the desired product is available, there is not likely to be a shift in service expectations or attitude toward the retail outlet.

If the product is not available from the retail outlet, negative disconfirmation toward the outlet is realized by the customer, leading to a level of dissatisfaction. The customer may experience some immediate change in their attitude toward loyalty and will contemplate a number of behavioral alternatives each of which, depending on the outcome, may contribute to a further change in attitude toward the primary retail outlet as well as a shift in expectations of service during the next service incident.

Components of the model draw from theories

associated with the determinants of customer expectations of service quality and satisfaction (Zeithaml, Berry, and Parasuraman 1996, Westbrook 1981, Oliver 1980, Dabholkar 1995a, Dabholkar 1995b, Dabholkar, Thorpe, and Rentz 1996), the role of disconfirmation as an antecedent to the customer assessment of dis/satisfaction with service encounters (Westbrook 1981, Rust and Zahorik 1993, Parasuraman, Zeithaml, and Berry 1985), the transition of a dis/satisfaction assessment into a more global attitude construct (Oliver 1980, Bolton and Drew 1991), the development of behavioral intentions as a result of the combined dis/satisfaction and attitude assessment (Zeithaml, Berry, and Parasuraman 1996), the decisions associated with purchase behavior involving out of stock conditions (Emmelhainz, Stock, and Emmelhainz 1991, Emmelhainz, Emmelhainz, and Stock 1991, Schary and Christopher 1979), and the post consumption assessment of product satisfaction (Oliver 1980, 1989, Tse and Wilton 1988, Woodruff, Cadotte, and Jenkins 1983, Spreng and Mackoy 1996).

Customer Satisfaction with Retail Outlet Service Quality

The first component of the model addresses the assessment of service dis/satisfaction and development of attitudes toward the service provider, in this case a retail outlet. The service transaction is initiated by the recognition of a product need on the part of the customer. Based on that need the customer visits the primary retail outlet in order to purchase the product. Depending on previous experience with the retail outlet the customer may have formed some basis of loyalty to patronize a specific retailer and store.

Customer Evaluation of Retail Service. The theoretical basis for both customer dis/satisfaction and service quality evaluations is the disconfirmation paradigm originally received from social psychology research (Oliver 1980). The disconfirmation paradigm involves the formation of expectations regarding product or service performance and the disconfirmation of those expectations as a result of some performance comparison. While the customer satisfaction construct is typically associated with the post

consumption evaluation of product performance versus a comparison standard, service quality is generally associated with an assessment of service attributes based on more idealized performance standards (Dabholkar 1993). Satisfaction and service quality have generally been supported by separate streams of research. While service quality research has focused on attributes associated with service performance, dis/satisfaction research has primarily focused on product evaluation. Research in satisfaction, however, is beginning to incorporate service phenomena. Recent debates suggest there are circumstances when measures of one or both constructs, (satisfaction or service quality) may provide better insights into customer evaluations of service (Dabholkar 1995a).

In a series of articles, Dabholkar (1993, 1995a, 1995b) investigated the similarities and differences between the satisfaction and service quality constructs. Her original analysis distinguished the constructs in the literature in that dis/satisfaction expectations are based on experience or another evaluative standard rather than the ideals which are used in service quality; dis/satisfaction takes a transaction level perspective evolving into a more global attitude while service quality is viewed as a more global concept; evaluations of dis/satisfaction contain aspects of affect as well as cognition while service quality is primarily cognitive in nature. Dabholkar (1993) speculated that whereas the two constructs were likely to overlap under certain conditions, there was potential for the constructs to be operationalized as separate factors and the possibility of a causal link between the two constructs. In a following article, Dabholkar (1995a) offered further theoretical evidence of a causal link which varied according to contingencies between the dis/satisfaction and service quality constructs. Dabholkar (1995b) confirmed her original speculation regarding construct similarities and distinctions with an empirical examination which identified customer dis/satisfaction and service quality as separate factors when viewed in the short term, and converging into one indistinguishable factor in the long term based on the cognitive and affective aspects of the constructs. This finding appears to have been confirmed by Spreng and Mackoy (1996) in an

empirical study of student assessments of undergraduate advising. Their study evaluated both dis/satisfaction and service quality operationalizing service quality based on a distinction between desires and evaluation of service received.

The representation of both satisfaction and service quality constructs are important to the temporal application of the model depicted in Figure 1. Viewing a customer transaction involving a stockout situation as an isolated incident suggests that the satisfaction construct may form the basis for comparison incorporating previous experience with the current or an alternate retail outlet. Over an extended relationship with a retail outlet involving a series of shopping experiences, either satisfaction or service quality play a role in the disconfirmation process.

A second debate more directly associated with consumer dis/satisfaction research concerns standards of comparison used in the evaluation of consumer dis/satisfaction. Woodruff, et al (1991) provide an assessment of alternative comparison standards presented in the dis/satisfaction literature, their distinctions and commonalities, and the implications of using different standards in a retail setting.

The model of customer reaction to stock availability (Figure 1) considers that expectations, as designated by the Service Expectations construct, are based on experienced-based norms originally conceptualized as part of a modified confirmation/disconfirmation process presented by Woodruff, Cadotte, and Jenkins (1983). Within the experience-based norms model, expectations are derived from a variety of personal experiences involving the personal use of a product, word-of-mouth endorsements and criticisms, as well as marketing efforts presented by companies. Woodruff, Cadotte, and Jenkins (1983) identified three different outputs resulting from this prior experience: expectations about the focal brand, experienced-based performance norms, and brand attitudes.

Service expectations at a retail outlet as represented in Figure 1 are proposed to originate from previous experiences with the retail outlet. In service industries, much as with products, customers may have prior service experiences with

the primary company as well as other companies offering similar services, word-of-mouth endorsements and criticisms may be presented by others, and the service provider may be involved in marketing efforts that contribute toward customer beliefs about services offered. An interesting component of the experience-based norms model is the incorporation of cognitive as well as affective processes. Woodruff, Cadotte, and Jenkins (1983) posited that brand attitude is influenced indirectly by prior experience through the mediating variables of expectations and performance norms. In the model presented in Figure 1, this would affect attitudes toward the retailer if they were viewed as a brand name service provider.

Westbrook (1981) distinguished customer dis/satisfaction with retail establishments from product/service dis/satisfaction suggesting that dis/satisfaction with a retail establishment can be viewed as:

“an individual’s emotional reaction to his or her evaluation of the total set of experiences realized from patronizing that retailer. In contrast, product/service satisfaction refers to the consumer’s emotional response to his or her evaluation of experience obtained from the usage, consumption, and ownership of the specific good or service.”

By this he incorporated a diverse set of factors associated with the retail experience into a more global assessment of dis/satisfaction and attitude toward the retail outlet. Westbrook’s model also recognized the potential for changes in retail dis/satisfaction as a function of changes in the components of the experience. Stock level is proposed in Figure 1 as one component which may affect customer assessment of dis/satisfaction toward the retail outlet.

Based on a series of focus group interviews with customers of various service firms, Zeithaml, Berry, and Parasuraman (1993) conceptualized a framework which incorporated a zone of tolerance regarding customer expectations. The zone of tolerance, conceptualized as a zone of indifference in previous work by Miller (1977), is bounded by two standards of expectation; “desired service,”

which they defined as, "the level of service a customer hopes to receive," and, "adequate service," which they defined as, "the level of service the customer will accept."

Zeithaml, Berry, and Parasuraman (1993), suggested that the zone of tolerance represented a recognition and acceptance by customers of the heterogeneous nature of services. As a result, if perceived service falls within this range it is unlikely that the customer will experience any form of disconfirmation toward the service offering. If perceived service does not fall within the zone of tolerance, the customer will experience disconfirmation toward the service and will form an evaluation of satisfaction or dissatisfaction, depending on the circumstances. Figure 1 presumes a similar zone of tolerance in the assessment of disconfirmation toward the service provided by the retail outlet. In this case, if the level of stock availability is not adequate to meet the needs of the customer, (the product is completely out-of-stock, fewer items are available in stock than are needed by the customer, or perhaps an adequate number of items are available in stock, however they appear to be in poor physical condition), negative disconfirmation toward the service provider will result. On the other hand, if adequate stock is available, customer expectations are confirmed.

In addition to recognizing a zone of tolerance associated with customer assessments of disconfirmation, the framework presented by Zeithaml, Berry, and Parasuraman (1993) recognized two other specific factors affecting the assessment of service disconfirmation which are considered in the model in Figure 1. First, if interpreted as an *implicit service promise*, stock availability represents an antecedent of both desired and predicted service. Second, *perceived service alternatives*, such as service provided by alternative retail outlets, represent an antecedent to the formation of a level of adequate service.

The model presented in Figure 1 proposes that attitude toward the retail outlet may also be influenced by previous experiences with the primary retail outlet as well as competing retail outlets. The model also suggests that, in circumstances in which alternative products must be purchased as a result of an out of stock position on the original item, the attitude toward the retail

outlet can be influenced by an assessment of product dis/satisfaction.

For example, should the desired product be available, the likely scenario will result in the confirmation of expectations. Depending on the circumstances of the need and the nature of the product, such as a highly sought toy at Christmas, the customer may very well experience a level of positive disconfirmation associated with the product availability.

If, however, a desired product is in an out-of-stock inventory position when the customer arrives at the shelf location, the stockout condition in the retail outlet results in an initial assessment of negative disconfirmation toward the outlet, which contributes to the customer's attitude toward service quality and intentions toward patronage.

Factors Affecting Customer Loyalty. Bolton and Drew (1991) pointed out that when frequently repeated transactions are evaluated, the distinction between satisfaction and a broader attitude toward the service provider becomes "blurred." Oliver (1980) applied adaptation level theory to assess the impact of post usage ratings on attitude and expanded the model to include future purchase intentions. His findings suggested that dis/satisfaction outcomes can influence post purchase attitude as well as future purchase intentions. While the model presented by Oliver is directed at product dis/satisfaction, its applicability to service breakdowns as suggested in Figure 1 may influence future patronage decisions in two ways. First, as a direct adaptation to current attitudes toward the original retail outlet and, second, in circumstances where the service breakdown leads to an experience with an alternative retail outlet, dis/satisfaction associated with that experience may also influence expectations toward the primary retail outlet.

Schary and Christopher (1979) noted customer reactions to stockouts may be affected by brand loyalty, store loyalty, and lack of expressed loyalty. In addition they recognized that while a stockout may weaken a customer's bond to a store, the stockout must first be observed by the customer. Such an observation may be affected by customer shopping behavior. Customers using a pre-planned shopping list are likely to observe a stockout while others, who may use the store as a

devise to remind them of product needs, may not realize the out of stock position.

Investigating the impact of service improvement and customer satisfaction, Bolton and Drew (1991, p.2) noted the following relationship between dis/satisfaction and attitude:

“Researchers distinguish between two constructs, customer satisfaction and attitude. Customer *satisfaction* refers to a customer’s evaluation of a *specific transaction*. In contrast, a customer’s *attitude* corresponds to a *global* evaluation of the product/service, rather than to an evaluation of a specific transaction.”

They presented a model to explain the impact of service dis/satisfaction on attitude. The equations presented recognized the formation of a feedback loop from the primary service provider that affects attitude and future purchase intentions based on the service dis/satisfaction. The model in Figure 1 recognizes such a feedback loop leading from customer purchase behaviors to the “Service Expectations” and “Attitude Toward the Retail Outlet” constructs.

Customer Reactions to Retail Stock Availability

Actions customers may take in response to variations in retail stock availability incorporate the development of behavioral intentions, followed by one of four alternative behaviors as indicated in Figure 1. If the desired product is in stock at the retail outlet at the time of the customer’s visit, the customer will buy the product. If the desired product is not in stock from the retail outlet at the time of the customer’s visit, the customer may decide not to purchase the product at the present time, may decide to purchase a substitute product considering alternative sizes or brands, or the customer may decide to visit an alternative retail outlet in the hope of obtaining the product originally desired.

Behavioral Intentions Regarding Stockout Conditions. Zeithaml, Berry, and Parasuraman (1996) recognized that changes in customer dis/satisfaction lead to behavioral intentions. The central intention addressed in their model involved

the potential for customer defection. Highlighting the implications of dissatisfaction on financial performance, their model hypothesized alternative types and levels of behavioral intentions based on the customer’s perception of service quality performance. Their model considered service quality outcomes where customers do and do not experience problems with service as well as when service problems are resolved. Examples of resulting behavioral intentions included favorable intentions such as saying positive things about the company, remaining loyal to the company, spending more money with the company, as well as unfavorable intentions such as negative word-of-mouth, switching to another company, reducing the amount of business conducted with the company, and so forth. Results of a study conducted with customers of four different types of service firms, including a retail chain, suggested that maintaining service quality within the zone of tolerance sustains the customer’s favorable intentions toward patronage. However, companies must reach the customers’ desired level of service quality if they wish to raise favorable behavioral intentions toward patronage.

Their findings also suggested that under conditions of service recovery it is unlikely that customer behavioral intentions will remain at previous levels. Within a retail circumstance, an example may be if an out of stock is experienced by a customer and the retail outlet arranges for a special delivery of the product or provides for an attractive discount on a demonstration model. Results of their study suggested that service failures may weaken the customer-company bond even with satisfactory problem resolution as described above.

Decisions and Behaviors Resulting From Out of Stock Circumstances. A number of researchers have investigated customer reactions to retail stockout conditions, focusing on the choice alternatives considered and the decision made when experiencing a stockout condition.

Schary and Christopher (1979) identified five different stockout behaviors including options to go to a different store, canceling the purchase, delaying purchase, buying an alternate brand, or buying an alternate size.

Emmelhainz, Stock, and Emmelhainz (1991a)

manipulated out of stock conditions on selected products in a grocery store and identified 15 different possible behavioral reactions to stockouts involving decisions whether or not to purchase at the store, to go elsewhere, or to select alternative brands, sizes, or varieties of product. An important finding of their study identified three factors which contributed to the decision whether or not to substitute items: perceived product risk, intended product usage, and urgency of need.

Walter and Grabner (1975) identified six potential reactions to a stockout condition at a liquor store. The six factors included decisions to return later, patronize another store, or substitute another item considering either size, brand, or cost.

The model of customer reaction to stock availability summarizes alternatives into the three resulting behaviors associated with a stockout condition. One option illustrated in Figure 1 is to buy a substitute item from the same retail outlet. This item may be the same brand in a different size or variety, a different brand and so forth. Under this scenario the model indicates that a subsequent dis/satisfaction process would be entered. This process would focus on an assessment of dis/satisfaction with the alternative product.

Rather than buy a substitute, the customer may choose to patronize an alternate retail outlet. Such a move may lead to the purchase of the desired product and confirmation regarding product performance, however the stockout circumstance at the original retail outlet creates an opportunity for customer exposure to service at the alternate outlet. As with the alternate product loop, it is proposed that the initial out of stock position contributes to a level of dissatisfaction with the original service provider, as well as modifying service expectations and maybe more importantly, a shift in customer attitude toward the service provider.

The third behavior is a decision not to buy the product at this time. This is proposed not to affect expectations or changes in attitude other than that associated with the initial level of disconfirmation.

Product Impact on Service Expectations and Attitude

While the model does not explicitly detail the disconfirmation process associated with an evaluation of product dis/satisfaction, the "Product Dis/satisfaction" construct represents a second area of the model where the disconfirmation paradigm is applied. In this case disconfirmation presents a means to evaluate product dis/satisfaction. This application of the dis/satisfaction model is more traditional and, as with the service disconfirmation illustrated in the top level of the model, this level of dis/satisfaction evaluation is also associated with the model of experience-based norms presented by Woodruff, Cadotte, and Jenkins (1983).

The role of the Product Dis/Satisfaction construct within the overall model is to suggest possible alternative outcomes based on the customer's assessment of dis/satisfaction associated with the product purchased. The model proposed that the outcome of the product evaluation may influence expectations and attitudes associated with the service provided by the retail outlet. For example, under circumstances where the product to be purchased is the original product desired by the customer, the outcome from the product dis/satisfaction model is likely to be an assessment of dis/satisfaction with only the product. If, however, the product is an alternate product than that originally desired, the outcome can have an extended impact on expectations of service quality as well as loyalty toward the retail outlet. Should the alternate product perform in a manner that results in positive disconfirmation, the possibility exists that the customer will attribute some of their good fortune to the retail outlet service. This attribution may affect the customer's general attitude toward the outlet but will likely not affect expectation toward stock availability. This affect is represented by the path leading from the "Product Dis/Satisfaction" construct to the "Attitude Toward Retail Outlet" construct.

If, on the other hand, the evaluation of product dis/satisfaction results in an assessment of poor performance compared to product expectations and an associated negative disconfirmation, there is the possibility that the customer will, in this case, attribute some of their misfortune to the retail outlet, leading to a decrease in overall loyalty and

a shift in service expectations.

To summarize the impact of alternative product selection, it is proposed that the dis/satisfaction experienced as a result of the alternate product may have a secondary effect on customer expectations and attitude toward the product. If the product experience results in a positive disconfirmation evaluation, the customer may in fact increase their satisfaction with, and loyalty toward, the retail outlet. A negative disconfirmation evaluation, on the other hand, may lead to dissatisfaction with, and a reduction in loyalty toward, the retail outlet.

Alternate Retail Outlet Impact on Service Expectations and Attitude

Alternate retail outlets offer customers a basis for comparison regarding service expectations (Zeithaml, Berry, and Parasuraman 1996), as well as an alternative source for desired products. Obviously, for the primary retail outlet, an alternate outlet represents a threat to future revenues.

Within the framework of the model, the level of dis/satisfaction with an alternate retailer helps to establish the base of "adequate service" within the construct of "Service Expectations" used in making a comparison judgment. As a result, any exposure to alternate retail outlets providing the same products and service presents a mechanism by which customers may alter expectation of the primary outlet. The process associated with the behavior of a customer going to an alternate retail outlet mirrors the overall model presented in Figure 1.

If the customer experiences positive disconfirmation with the shopping experience at the alternate outlet, it is proposed that the experience will trigger a secondary assessment which may contribute to a negative attitude toward the primary outlet as a result of the original stockout. This affect is represented by the path leading from the "Alternate Retail Outlet" construct to the "Attitude Toward Retail Outlet" construct. On the other hand, an alternative may, in fact, enhance the assessment of service provided by the primary retail outlet. Under one possible scenario, the secondary retail outlet is also out of the item, suggesting a general shortage and

reduction in dissatisfaction with the primary retail outlet.

IMPLICATIONS FOR RESEARCH AND PRACTICE

The model of customer reactions to stock availability developed in this paper offers an explanation of the role that retail inventories can play in supporting customer satisfaction and retention. The model combines theoretical perspectives associated with customer assessments of service quality, satisfaction, and attitude toward the retail organization, along with the development of intentions and behaviors as a result of stock conditions, and the assessments of satisfaction with alternative products and service providers. This paper is intended to fill a gap in the literature by considering not only the immediate impact of stock availability on the customer's satisfaction with the retail outlet, but, in cases when a stockout situation is realized, the potential for a further assessment of satisfaction toward the retail outlet based on the performance of a substitute product or experience with an alternate retailer.

For practitioners, the model provides a conceptual framework to consider a broader set of implications resulting from stocking policies in the retail environment. Previous research regarding stock conditions has centered on immediate behaviors exhibited by customers who experience a retail stockout condition. By understanding the role of stock conditions in the development of attitude toward the retail outlet, merchandising, marketing, and inventory managers can better establish a cost of service to incorporate into alternative marketing and merchandising strategies. If, for example, a relationship between stock availability and customer attitudes toward patronage can be quantified as suggested by the results from the Schary and Christopher study, the model of customer reactions to stock availability offers a means to establish more effective in-stock service objectives that optimize the cost of holding retail inventories, the cost of immediate lost sales from out-of-stock inventory positions, as well as the potential cost of future lost sales due to lost patronage.

Rather than considering inventory manipulation as a sole means to control customer

service, retail organizations may be able to address service issues by more effectively managing customer expectations of stock availability in the development of their marketing and service strategies. Such expectations may consider factors such as an assessment of stock availability provided by competitors, the availability of alternative products, and the messages presented to current and potential customers through various forms of communication.

Stock availability includes breadth as well as depth, and by understanding customer behavior toward stock selection and loyalty, merchandise managers will be able to consider short and longer term implications of adding and deleting product brands, variety, and sizes.

From a research perspective, a number of opportunities for future investigation regarding the impact of stock availability on customer satisfaction and retention can be identified. An important next step will begin to incorporate methods which will be useful in measuring the relationships identified in the model. The ability to establish and apply objective measures of service and satisfaction has been demonstrated by Rust and Zahorik's (1993) approach in assessing the contribution of attributes of customer dis/satisfaction to improved retention, market share, and profitability in the banking industry. The retail context considered in this model adds some complexity to the development of measures because of the relative freedom retail customers have for brand switching compared to that found in the banking industry.

The role of stock availability as a basis for customer expectations regarding service is not clearly understood. If merchandising is incorporated into the set of store characteristics which impact retail customer satisfaction (Westbrook 1981), how much of that impact can be attributed to stock availability based on inventory levels rather than breadth of selection?

As stated previously, the model presented in Figure 1 applies the concept of experienced based norms in the establishment of service expectations (Woodruff, Cadotte, and Jenkins 1983). Within the retail environment, further investigation is needed to assess the influence of various sources on the development of such norms. Potential sources include direct experience with the retail

outlet as well as mechanisms used by the retailer to define service capabilities for customers, and experience with competing retail outlets in establishing expectations for stock availability.

The model also points to an opportunity to investigate links involving secondary feedback loops between customer satisfaction with alternative retail outlet service or customer satisfaction with a substitute product, and its influence on attitude toward the primary retail outlet.

Rather than focusing on research questions associated with components of the model, additional research may address the applicability of the model under different retail conditions. For instance, it would be interesting to investigate the applicability of the model in a specialty goods retail environment. The nature of the customer clientele may be different as well as their behavior toward stockout conditions. Hansen and Deutscher (1977) in their investigation of retail store image found no significant difference in the measure of important attributes between grocery and department store image. This suggests that while the application of this model is centered on general merchandise retailers, it may also be applicable under specialized merchandise retail scenarios. However, a distinguishing factor which may prevail in a specialty goods retail environment concerns the role of product involvement in the assessment of retail satisfaction. Does the level of product involvement affect service expectations and attitude toward the retail outlet when stockout conditions are experienced by the customer? Higher involvement specialty products may limit the number of substitute products or alternate retail outlets.

Finally, while the model is developed based on stock availability as a central phenomenon, its applicability may extend to other individual or groups of measures associated with evaluating customer reactions to service criteria identified with the retail environments.

REFERENCES

- Andersen Consulting (1996), *The Retail Problem of Out-of-Stock Merchandise*, Coca Cola Retailing Research Council.
- Ballou, Ronald H. (1992), *Business Logistics Management*, Third Edition, Prentice-Hall Inc., Englewood Cliffs,

- NJ.
- Bolton, Ruth N. and James H. Drew (1991), "A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes," *Journal of Marketing*, 54, (April), 69-82.
- Churchill, Gilbert A. Jr. and Carol Surprenant (1982), "An Investigation Into the Determinants of Customer Satisfaction," *Journal of Marketing Research*, 19, (November), 491-504.
- Dabholkar, Pratibha A. (1993), "Customer Satisfaction and Service Quality: Two Constructs or One?" *American Marketing Association Educators' Proceedings*, David W. Cravens and Peter R. Dickson, (eds.), 4, 32-43.
- Dabholkar, Pratibha A. (1995a), "A Contingency Framework for Predicting Causality Between Customer Satisfaction and Service Quality," *Advances in Consumer Research*, 22, 101-108.
- Dabholkar, Pratibha A. (1995b), "The Convergence of Customer Satisfaction and Service Quality Evaluations with Increasing Customer Patronage," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 8, 32-43.
- Dabholkar, Pratibha A., Dayle I. Thorpe and Joseph O. Rentz (1996), "A Measure of Service Quality for Retail Stores: Scale Development and Validation," *Journal of the Academy of Marketing Science*, 24, (1), 3-16.
- Emmelhainz, Margaret A., James R. Stock and Larry W. Emmelhainz (1991), "Consumer Responses to Retail Stock-outs," *Journal of Retailing*, 67, (2), 138-147.
- Emmelhainz, Larry W., Margaret A. Emmelhainz and James R. Stock (1991), "Logistics Implications of Retail Stockouts," *Journal of Business Logistics*, 12, (2), 129-142.
- Hansen, Robert A. and Terry Deutscher (1977), "An Empirical Investigation of Attribute Importance in Retail Store Selection," *Journal of Retailing*, 53, (Winter), 59-72.
- Miller, John A. (1977), "Studying Satisfaction, Modifying Models, Eliciting Expectations, Posing Problems, and Making Meaningful Measurements," *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*, H. Keith Hunt, (ed.), 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*, 27, (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.
- Parasuraman, A., Valarie A. Zeithaml and Leonard L. Berry (1985), "A Conceptual Model of Service Quality and its Implications for Future Research," *Journal of Marketing*, 49, (Fall), 41-50.
- Rust, Roland T. and Anthony J. Zahorik (1993), "Customer Satisfaction, Customer Retention, and Market Share," *Journal of Retailing*, 69, (2), 193-215.
- Schary, Philip B. and Martin Christopher (1979), "The Anatomy of a Stock-Out," *Journal of Retailing*, 55, (2), 59-70.
- Spreng, Richard A. and Robert D. Mackoy (1996), "An Empirical Examination of a Model of Perceived Service Quality and Satisfaction," *Journal of Retailing*, 72 (2), 201-214.
- Tse, David K. and Peter C. Wilton (1988), "Models of Consumer Satisfaction Formation: An Extension," *Journal of Marketing Research*, 25, May, 204-12.
- Walter, C. K. and John R. Grabner (1975), "Stockout Cost Model: Empirical Tests in a Retail Situation," *Journal of Marketing*, 39, (July), 56-60.
- Walter, C. K. and Bernard J. LaLonde (1975), "Development and Tests of Two Stockout Cost Models," *International Journal of Physical Distribution*, 5, (3), 121-132.
- Westbrook, Robert A. (1981), "Sources of Consumer Satisfaction with Retail Outlets," *Journal of Retailing*, 57, (3), 68-85.
- Woodruff, Robert B., Ernest R. Cadotte and Roger L. Jenkins (1983), "Modeling Consumer Satisfaction Processes Using Experienced-Based Norms," *Journal of Marketing Research*, 20, (August), 296-304.
- Woodruff, Robert D., 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.
- Zeithaml, Valarie A., Leonard L. Berry and A. Parasuraman (1993), "The Nature and Determinants of Customer Expectations of Service," *Journal of the Academy of Marketing Science*, 21, (1), 1-12.
- Zeithaml, Valarie A., Leonard L. Berry and A. Parasuraman (1996), "The Behavioral Consequences of Service Quality," *Journal of the Academy of Marketing Science*, 60, (2), 31-46.

Send correspondence regarding this article to:

Carlo D. Smith, Doctoral Student
 Department of Marketing, Logistics, and Transportation
 The University of Tennessee
 310 Stokely Management Center
 Knoxville, TN 37996-0530 USA

TOWARDS A COMPREHENSIVE TAXONOMY AND MODEL OF CONSUMER COMPLAINING BEHAVIOUR

Jonathan Boote, University of Luton

ABSTRACT

The most widely used taxonomy of consumer complaining behaviour (Singh 1988) is limited in two respects: (1) it oversimplifies the key types of complaining behaviour - for example negative word-of-mouth need not just be a private action; and (2) it fails to appreciate that complaining behaviour often has two stages as certain CCB types (such as third party action) may only be entered into once other CCB types have failed to generate a satisfactory level of perceived justice. This paper offers a two-factor taxonomy of CCB which takes into account these issues. Complaint types are classified in terms of whether they are primary or secondary, and in terms of whether they are involved or uninvolved. This paper also offers a comprehensive model of CCB, which builds on the conceptual approach of Blodgett and Granbois (1992) by considering the whole CCB process as having four stages: (1) cognitive reasoning; (2) affective response; (3) triggers of consumer dissatisfaction responses; and (4) affective action. It is argued that there are eight major triggers of which lead to a certain dissatisfaction response (or a set of responses); each trigger being made up of a number of dimensions. The paper concludes by discussing possible directions for future research.

INTRODUCTION

Singh (1988) conceptualised the phenomenon of consumer complaining behaviour (CCB) as "a set of multiple (behavioural and nonbehavioural) responses, some or all of which are triggered by perceived dissatisfaction with a purchase episode" (p94). Work in the area of CCB can be divided into three broad areas: (1) the development and testing of theories of consumer dissatisfaction - which provide the theoretical starting point for complaining behaviour; (2) the study of complaining behaviour types, out of which taxonomies, typologies and models of complaining behaviour have been developed; and (3) the analysis of various triggers of complaining behaviour - which move consumers from the

affective response of consumer dissatisfaction to the affective action of complaining behaviour. However, as East (1998a) argues, much of the work conducted on CCB triggers has been piecemeal and "a method is required that covers all the potential causes of complaining so that the relative influence of different factors can be established" (p.401).

Indeed, from a classical Kuhnian perspective, CCB research is still in its infancy (or pre-paradigm stage) where basic classifications, models and approaches are still being debated; and where much empirical evidence appears to contradict earlier research - especially findings relating to CCB triggers. In a recent attempt to address this problem, East (1998a) proposed the theory of planned behaviour (Ajzen 1991) as an all-inclusive theoretical tool for the analysis of the bases of consumer complaining. However, the theory does not appear to have worked in practical scenario-driven experiments (East 1998b) and it also neglects certain potential triggers of CCB such as demographics (which, admittedly, have low predictive powers) and perceptions of attribution. Acknowledging the disparate nature of the discipline and the apparent failure to fit an inclusive theory around CCB, this paper consolidates previous theoretical approaches to complaining behaviour through the development of a comprehensive taxonomy and model of CCB - both of which may be used as a basis for future empirical research.

Note that in the discussion of complaining that follows, the analysis assumes a strong link between dissatisfaction and CCB. The starting point for this paper is dissatisfaction with a product or service. Other researchers (such as Kowalski, 1996) rightly argue that some consumers complain not out of dissatisfaction, but in an effort to win concessions from a retailer or manufacturer. However, such consumers are outside the scope of this paper, which focuses on the *genuinely* dissatisfied.

CONSUMER DISSATISFACTION: A SUMMARY OF THEORETICAL APPROACHES

The starting point for most models of CCB is consumer dissatisfaction. However, the fact that no one theory of consumer dissatisfaction is universally accepted by academics in the field of complaining behaviour is due to the problem of 'standards' (Woodruff et al. 1991). Dissatisfaction is usually conceptualised as the outcome of a comparison to a standard, although this theory is now being questioned (Yi 1990). If one accepts the comparative approach to dissatisfaction, then a key question in CCB research is: which standard(s) do consumers use to evaluate a purchase? Combining the work of Woodruff et al. (1991) and Erevelles and Leavitt (1992), there are at least six possible theoretical approaches to dissatisfaction: disconfirmation of expectations, attribution, equity, experienced-based norms, perceived performance, response to an ideal, comparison based on promises, and a non-cognitive, affective approach. Of all these theories, the three most widely discussed in the complaining behaviour literature are disconfirmation of expectations, attribution and equity. However, the question still remains as to which theory best works in practice, as Everelles and Leavitt (1992) maintain that consumers may well apply (simultaneously) different standards of comparison.

Disconfirmation of Expectations and Attribution Theory

The most widely accepted theory of consumer dissatisfaction is that of disconfirmation of expectations. If perceived quality is lower than expectations, then negative disconfirmation is said to be the resultant cognitive state, with consumer dissatisfaction conceptualised as the resultant affective state. Disconfirmation, as an all-embracing theory of consumer dissatisfaction, has come under criticism (Erevelles and Leavitt 1992), as it has been argued that disconfirmation, in all circumstances, may not be enough to cause dissatisfaction. Folkes and Kotsos (1986) argue that a consumer's perception of the attribution of product or service failure will moderate feelings of dissatisfaction. If the cause of disconfirmation is

externally attributed (i.e. not caused by the consumer), then the consumer is justified in feeling dissatisfied. However, if the disconfirmation is internally caused (i.e. the fault of the consumer) - for example, if instructions were not followed - then dissatisfaction ought not to be directed at the retailer, or the manufacturer, involved. In other words, if negative disconfirmation is externally attributed, a consumer is not justified in engaging in complaining behaviour.

Equity Theory

An alternative perspective of consumer dissatisfaction is provided by proponents of equity theory - such as Tse (1990) and Lapidus and Pinkerton (1995). Equity theory is concerned with the balance, and perceived fairness, of the inputs and outputs of a particular transaction. From the perspective of either side, there are three possible outcomes of a given transaction as prescribed by equity theory: (1) equity, (2) positive inequity, (3) negative inequity. Equity is the case where inputs and outputs of either side are perceived to be of an equal degree. Inequity exists where one side in the transaction is perceived to have gained the upper hand. Positive inequity is the case where, from your point of view, you have gained more from the transaction, either in terms of inputs or outputs, than the other side. Negative inequity is the case where the other side is perceived to have gained more than you. From an equity perspective, consumer dissatisfaction is the result of negative inequity, where the consumer perceives to have gained less from a transaction than the seller. A complaining behaviour is, therefore, likely if dissatisfaction is caused by negative inequity.

Alternative Approaches to Dissatisfaction

Experienced-Based Norms. A further standard used as a reference point in the interpretation of consumer dissatisfaction is experienced-based norms (see Woodruff et al. 1983). Disconfirmation is said to be the result of a comparison of the most current purchase with a past purchase - either of the same brand or a different brand in the same product class. In an empirical study by Cadotte et al. (1987), both the product-based and the brand-based norms were

considered stronger explanations of consumer satisfaction than the disconfirmation-of-expectations approach.

Comparison to an Ideal. A question that is increasingly raised in the consumer dissatisfaction literature is: how realistic are consumers in forming their expectations? This issue was first raised by Miller (1977) in his categorisation of expectation standards. Do consumers expect a product to perform to a minimum tolerable standard, or to an 'adequate' level or to its apotheosis? The degree of expectation will obviously have an impact on the degree of (dis)satisfaction felt. Consumer dissatisfaction is increasingly probable the more the standard of expectations moves from the minimum tolerable level to the ideal.

Comparison to Promises Made by the Seller. Woodruff et al. (1991) argue that dissatisfaction may result from a disparity between what a seller promises (in terms of advertising, personal selling, packaging etc.) and the perceived quality of the purchase.

Perceived Performance. It is argued by some that, in reality, the rational, cognitive approach to consumer dissatisfaction implicit in the disconfirmation, attribution and equity interpretations do not hold true (Churchill and Surprenant 1982). Instead, it is asserted that (dis)satisfaction is caused simply by the perceived performance of the product or service irrespective of prior expectations. The idea that (dis)satisfaction is an affective response to the perceived 'goodness' or 'badness' of the purchase was also supported by the findings of Tse and Wilton (1988).

Affective/Emotional Approach to Consumer Dissatisfaction. There is a school of thought which asserts that cognitive elements of consumer dissatisfaction are overly stressed in the literature (Yi 1990). It has been argued that dissatisfaction is an affective or emotional state which can, or perhaps even does, bypass any cognitive process of evaluation. The implicit assumption in the cognitive approaches to dissatisfaction (such as disconfirmation of expectations) is that if dissatisfaction occurs, consumers will know

precisely what caused it - because a cognitive process of evaluation has been undertaken before the affective state of dissatisfaction is reached. What these cognitive approaches ignore is that consumers may feel dissatisfied without knowing the precise reasons why. In other words, a negative affective response to a purchase may come before a cognitive evaluation - especially if a further purchase is required in the future.

TYPES OF CONSUMER COMPLAINING BEHAVIOUR

If consumer dissatisfaction can be defined as an affective response to some form of negative cognitive reasoning following a purchase (or a purchase situation), then complaining behaviour can be interpreted as affective action. Hirschman (1970), widely regarded as the founder of the study of consumer complaining behaviour, argued that there are three possible responses to a worsening of quality in firms, organisations and states: exit, voice and loyalty. Although loyalty - meaning taking no action and remaining with the firm - is not often discussed in the complaining behaviour literature, exit and voice are well established as two of the four cornerstones of complaining behaviour along with negative word-of-mouth and third party action (Singh 1988). However, recent thinking suggests that it is oversimplistic to conceptualise complaining behaviour as a four-dimensional phenomenon (Huefner and Hunt 1994). Retaliation, grudgeholding and avoidance have also been discussed as complaining behaviours in their own right (Hunt and Hunt 1990).

Exit, Voice, Negative Word of Mouth and Third Party Action

The four most widely discussed complaining behaviours are exit, voice, negative word-of-mouth and third party action. Exit refers to a consumer who decides not to buy a product or service again, not to shop at a particular retailer or not to buy from a particular manufacturer again (or some combination of the above). Voice is an attempt to seek redress from the retailer or manufacturer involved (which can be either written or oral). The communication of dissatisfaction to family and

friends is classified as negative word-of-mouth - which is often in the form of a warning not to buy a certain product or to buy from a certain outlet. Third party action is the act of involving an outside agency to deal with a dissatisfying episode - such as a consumer group or a legal representative.

Retaliation, Avoidance and Grudgeholding

Huefner and Hunt (1994) put forward three further consumer complaining behaviours: retaliation, avoidance and grudgeholding. Retaliation is the process of 'getting even' with the seller; a form of revenge. Possible manifestations of retaliation, put forward by the authors, include destruction of products and equipment, theft, negative word-of-mouth in the store itself and disruption (such as putting items in the store in the wrong place). Both avoidance and grudgeholding are forms of extended exit. The problem with the current conceptualisation of exit is that it has no time-frame attached. Different consumers may exit for different lengths of time: some may return to buying the product after a week and some may never buy the product again. According to Huefner and Hunt, exit is a short-term phenomenon, whereas avoidance is more medium-term in a deliberate attempt to 'punish' the firm. Grudgeholding is much more extreme and can last years, if not decades.

A further problem with the term 'exit' (which becomes apparent during empirical research) is that there are four types of exit. A dissatisfied consumer can stop buying the brand, or can stop buying a particular product type (regardless of producer), or can stop buying from a particular retailer or manufacturer. These types of exit can develop over time into avoidance and grudgeholding. More empirical research is needed to establish the boundaries between types of extended exit. Exit-retailer and exit-manufacturer are more intense forms of exit than exit-brand as they involve a boycott of an entire range of products, not just one particular brand.

TAXONOMIES OF CONSUMER COMPLAINING BEHAVIOUR

Of all the numerous attempts to classify the

various types of CCB discussed in section 3 - see, for example, Day (1980); Bearden and Teel (1983); and Singh (1988) - the most accepted in the literature appears to be that of the latter. Singh (1988) took three of the key dimensions of complaining behaviour - voice, negative word-of-mouth and third party action - and classified them in terms of two dichotomies based on the object toward which the complaining behaviours are directed: internal/external and involved/uninvolved. The internal/external construct refers to whether or not the complaining behaviour is directed towards the dissatisfied consumer's social circle (i.e. internally directed) such as negative word-of-mouth, or directed outside the social circle (i.e. externally directed) such as voice. The involved/uninvolved construct is concerned with whether the object towards which the complaining behaviour is directed is involved in the dissatisfying experience. Voice would be classed as involved, whereas third party action would be considered uninvolved - because, for example, a legal representative did not directly cause the dissatisfying episode.

Out of this two-factor analysis of complaining behaviour, Singh developed a three-dimensional taxonomy. Singh tentatively argued that the complaining behaviour of 'no action' (which is referred to by Hirschman (1970) as 'loyalty') should be treated as a 'voice response' because - according to Singh's rationalisation - taking no action appears "to reflect feelings toward the seller" (p104). Singh's taxonomy is presented in Figure 1.

Perceived Justice and The Primary/Secondary Approach to Complaining Behaviour

The taxonomy given in Figure 1 does not accurately reflect many current ideas in consumer complaining behaviour. It is argued by many academics that complaining behaviour is a sequential process. Exit, negative word of mouth and, especially, third party action may only be entered into after voice has been used, and in circumstances when the consumer has not received a satisfactory level of 'perceived justice' (Blodgett and Granbois, 1992). Negative word-of-mouth, third party action and exit may only be used, therefore, if voice has failed. This sequential

dimension presupposes a primary/secondary classification of complaining behaviour whereby, using the traditional four types of CCB, voice is seen as a primary behaviour, negative word-of-mouth and exit may be either primary or secondary, and third party action is a secondary CCB. One could also classify retaliation, avoidance and grudgeholding as secondary complaining behaviours.

Figure 1
Singh's (1988) Taxonomy of Consumer Complaining Behaviour

Voice Responses	Private Responses	Third Party Action
Redress seeking (voice)	Warn family and friends (negative word of mouth)	Legal action
No action (loyalty)	Exit	Complain to consumer organisation

Perceived justice is an important concept in complaining behaviour research, as it is a moderator: it represents a standard by which a voiced complaint is assessed by the dissatisfied consumer (Blodgett and Granbois 1992; Blodgett and Tax 1993; and Blodgett 1994). As in the case with the analysis of dissatisfaction, the examination of perceived justice by a dissatisfied consumer is based around the concept of disconfirmation of expectations. If redress exceeds expectations, the result is positive disconfirmation and an affective state of perceived justice. If the consumer is satisfied with the level of perceived justice, then the dissatisfying episode is said to be closed, and, therefore, secondary complaining behaviours such as third party action and retaliation are not embarked upon. If redress falls short of expectations, then negative disconfirmation is the resultant cognition and perceived *injustice* is the resultant affective state.

There are three dimensions of perceived justice, any of which can contribute to a consumer's perception of (dis)satisfaction with the outcome of a dissatisfying episode: distributive,

procedural and interactional (Goodwin and Ross 1990; Blodgett and Tax 1993; and Blodgett 1994). Distributive justice is the perceived fairness of the tangible outcome. Interactive justice refers to the quality or fairness of interpersonal treatment during the conflict resolution stage. Finally, procedural justice refers to the perceived fairness of procedures and criteria used by decision-makers during conflict resolution. The dissatisfied consumer needs to feel that the decision reached is impartial and unbiased.

The Impact of Perceived Justice on CCB Classification

If we accept the primary/secondary approach to CCB types, and the central role that voice plays, this impacts upon how the phenomena are to be classified. It is, however, as discussed above, a distortion of reality to simply suggest that voice comes first, and all other CCB types are dependent on perceptions of justice relating to it. This is because: (1) other CCB types may be engaged in concurrently with voice; and (2) other CCB types may be used instead of voice. Therefore, it seems essential, in taxonomical terms, to sub-divide negative word-of-mouth and exit in relation to whether they occurred before (or alongside), or after, a voiced complaint. Separated by the concept of a 'redress boundary' (i.e. perceived justice arising from voice) are *pre-redress* negative word-of-mouth and *pre-redress* exit on the one hand, and *post-redress* negative word-of-mouth and *post-redress* exit on the other.

Third party action, retaliation, avoidance and grudgeholding are considered as solely secondary (i.e. post-redress) actions as they are most likely to occur as a result of a low level of perceived justice. Also included as secondary CCBs are voice, public negative word-of-mouth and post-redress exit behaviours.

Voicing may occur more than once. A reply from a firm may result in a low level of perceived justice, thus prompting the dissatisfied consumer to voice again, usually to an employee higher in rank.

Negative word-of-mouth is divided in terms of whether it is private (i.e. directed towards

people within the dissatisfied consumer's social circle) or public (i.e. directed to people outside the consumer's social network). Negative word-of-mouth is often private in the first instance - where the dissatisfaction is communicated just to close family relations and to friends. However, if the redress is not considered just, then negative word-of-mouth may become public - for example, by writing to a newspaper.

Exit behaviours considered as solely secondary CCB types (i.e. post-voice) are exit-retailer, and exit-manufacturer. Exit-brand and exit-product-category may be both primary and secondary CCBs.

A Two-Factor Taxonomy of Consumer Complaining Behaviour

As well as dividing CCB into primary and secondary actions, the taxonomy below includes the involved/uninvolved basis of classification as used by Singh (1988). The three involved CCB types (where the dissatisfied consumer has direct

contact with the firm) are primary and secondary voiced complaints, and retaliation. All other complaining types are considered uninvolved (where the dissatisfied consumer has either no, or indirect, contact with the firm). The taxonomy is presented in diagrammatic form in Figure 2. Note that avoidance and grudgeholding are shown in parentheses because they are extended forms of the four types of exit behaviour.

Note also that the taxonomy uses the term 'consumer dissatisfaction responses' (CDRs) rather than 'complaining behaviour' so that 'no action' and 'no further action' can be included in the classification as responses to dissatisfaction in their own right. The problem with previous classification attempts was that the complaining behaviour type of 'no action' seemed to fit somewhat artificially into the categorisation. In Singh's (1988) taxonomy, no action was included within voiced responses. The problem stems from having to consider no action as a behavioural rather than a non-behavioural response. In the introduction to this paper, complaining behaviour is referred to a "set of multiple behavioural and non-behavioural responses [to] ...dissatisfaction"

Figure 2
A Two-Factor Taxonomy of Consumer Dissatisfaction Responses

		Type	
		Uninvolved	Involved
Level	Primary	1. Exit-brand/product category 2. Private negative word of mouth 3. No Action	1. Primary voiced complaint
	Secondary	1. Post-redress exit behaviour (brand/product category/retailer/manufacturer) (2. Avoidance) (3. Grudgeholding) 4. Post-redress private negative word of mouth 5. Public negative word of mouth 6. Third party action 7. No further action	1. Secondary voiced complaint 2. Retaliation

(Singh, 1988 p.94). This definition is, in a sense, contradictory as a behaviour is said to include *non-behavioural* responses (such as taking no action). In order to get around this problem, it is suggested that the term 'complaining behaviour' be replaced with the term 'consumer dissatisfaction responses'.

THE TRIGGERS OF CONSUMER DISSATISFACTION RESPONSES

What, therefore, triggers how consumers react to dissatisfaction? This question has attracted a great deal of academic attention in recent years, where researchers have attempted to isolate one factor (or a number of factors) which affect how consumers react. Dissatisfaction is not thought to be a sufficient trigger by itself to cause a complaint, because, as studies have shown in both the US (Andreasen and Best 1977, TARP 1979) and the UK (Office of Fair Trading 1986), only about one in five dissatisfied consumers actually complain to the organisation concerned. Why is this the case? Recently, Kowalski (1996) conceptualised the issue in terms of thresholds: consumers have both a dissatisfaction and a complaining threshold. Consumers may be dissatisfied easily (i.e. they have a low dissatisfaction threshold) but may be reluctant to complain because they have a high complaining threshold. As the TARP and Office of Fair Trading studies demonstrate, there must be factors at work which act as a barrier to voiced complaining behaviour; factors which, it can be argued, trigger non-voiced complaining behaviour (i.e. exit and/or private negative word-of-mouth, or no action) These triggers, when taken together, will influence a consumer's complaining threshold - which may well vary with different dissatisfying experiences. Through a literature review, eight triggers have been identified; each being made up of a number of dimensions. These triggers can be seen as an extension of Andreasen's (1988) theory that CCB is caused by some interaction of four sets of factors: costs and benefits, personality, learning and restraints. Whilst it is beyond the scope of this paper to discuss in detail the research findings pertaining to each trigger, the dimensions of each trigger will briefly be considered below.

Situation

The situational triggers of CDRs refer to the specifics of the dissatisfying episode. Those situational triggers which have been previously examined include: product/service importance (Blodgett and Granbois 1992); level of involvement (Godwin et al. 1995); dissatisfaction intensity, (Prakash 1991); perceived costs and benefits of engaging in a particular CDR (Singh and Wilkes 1996); product/service cost (Kolodinsky 1993); product/service type (Singh 1990); and the practical causes of dissatisfaction - such as product recalls, service delays, and specific product characteristics (Standop 1991; Feinberg et al. 1996).

Attribution

There are two dimensions of attribution theory which are considered triggers of CDRs: perceptions of controllability and stability (Blodgett and Granbois 1992; and Singh and Wilkes 1996). Controllability refers to whether or not the dissatisfied consumer perceives that the company involved could have prevented the dissatisfying episode from occurring, and stability refers to the dissatisfied consumer's perception of whether the product/service failure is short or long term.

Demographics

Demographic factors linked to propensity to complain include age (Fails and Francis 1996); gender (Parker et al. 1993); income (Fails and Francis 1996); educational level (Kolodinsky and Aleong 1990); rural/urban location of dissatisfied consumer (Liefeld 1980); impact of having young children (Kolodinsky 1993); and the cost-sensitivity of the dissatisfied consumer (Parker et al. 1993).

Psychographics

Psychographic triggers of CDRs which have been examined include assertiveness/level of confidence and aggression (Richins 1983); attitude to, and past experience of, complaining, (Singh and Wilkes 1996); willingness to engage in uncomfortable situations (Tesser and Rosen 1975);

level of consumerism (Slama et al. 1993); personal values (Rogers et al. 1992); and locus of control (i.e. a belief in fatalism) (Foxman et al. 1990).

Company/Consumer Relationship

This trigger relates to such factors as degree of loyalty felt by the dissatisfied consumer to the company (Blodgett and Granbois 1992); company size (Kolodinsky and Aleong 1990); and the degree of interaction between company and consumer (Fornell and Didow 1980).

Marketplace/Consumer Relationship

It is argued that the market structure within which the company involved in the dissatisfying episode operates will have an effect on whether or not a consumer voices after a dissatisfying episode (Singh and Wilkes 1996).

Cultural Factors

This trigger is concerned with the issue of nationality: do consumers in some countries have

a higher propensity to complain than consumers in other countries? (Andreasen and Best 1977, Raven and Foxman 1994).

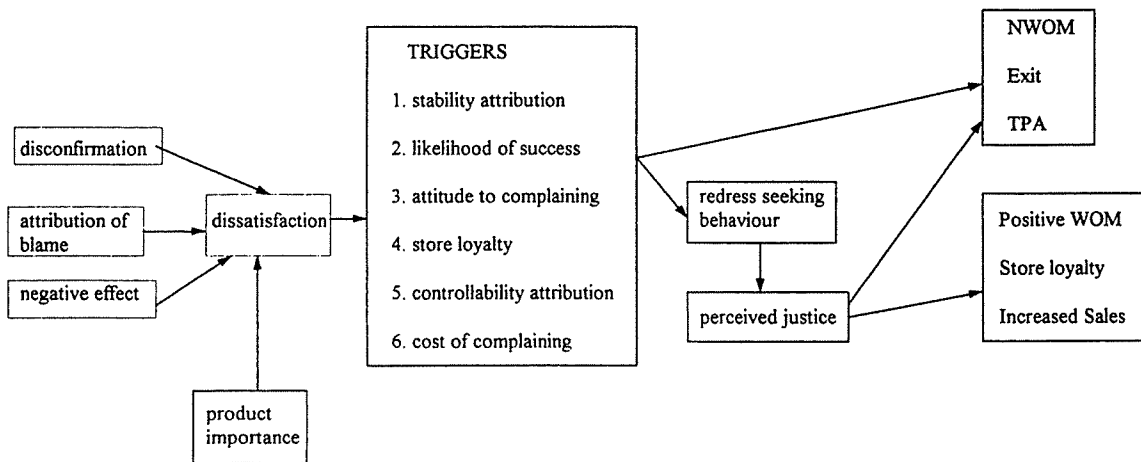
Social Factors

This trigger is concerned with the influence or persuasion of other people (i.e. the degree of responsiveness to peer pressure) (Malafi et al. 1993; and Slama and Celuch 1994).

PUTTING IT ALL TOGETHER: MODELS OF THE COMPLAINING BEHAVIOUR PROCESS

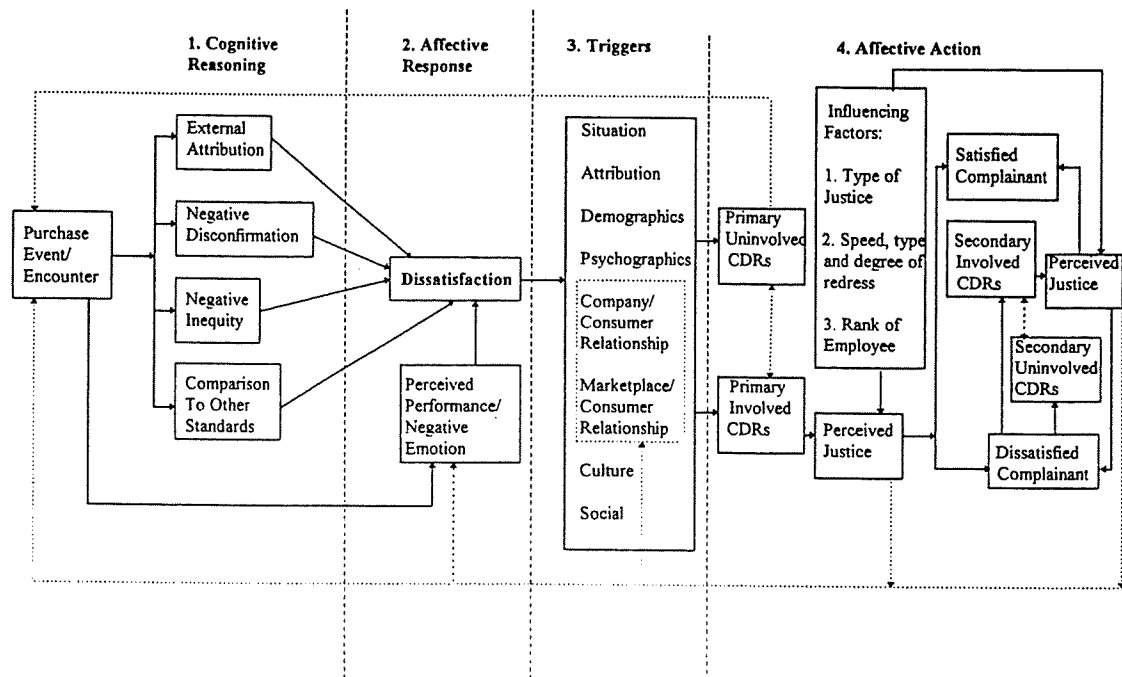
A number of theoretical models of complaining behaviour have been developed which seek to integrate work on both dissatisfaction, and taxonomies and triggers of complaining behaviour (see for example Day 1984; Nantel 1985; Blodgett and Granbois 1992; and Singh and Wilkes 1991). Of all these models, that of Blodgett and Granbois is the most comprehensive, and is shown in Figure 3.

Figure 3
Blodgett and Granbois' (1992) Conceptual Model of Consumer Complaining Behaviour



NWOM = negative word of mouth
WOM = word of mouth
TPA = third party action

Figure 4
A Four-Stage Conceptual Model of CONsumer Dissatisfaction Responses



There are three drawbacks to Blodgett and Granbois' model. Firstly, the complaining behaviour types of retaliation, grudgeholding and avoidance are not included. Secondly, the model does not include all possible complaining behaviour triggers such as demographic influences (e.g. age or gender), psychographic influences (such as aggression and attitude to complaining), alienation from the marketplace, and cultural and social influences. Thirdly, the model does not include all possible theoretical approaches to dissatisfaction: it only includes disconfirmation and attribution, together with a somewhat vague concept the authors call "negative effect".

In the light of these comments, a further model of complaining behaviour is now proposed which includes all the triggers of CDRs discussed earlier, as well as all possible responses to dissatisfaction which have been discussed in the complaining behaviour literature. The model uses the taxonomical approach shown in Figure 2, which classes complaining behaviour types in terms of two dichotomies: involved/uninvolved and primary/secondary. Note that all the triggers are treated equally - i.e. none are considered as

moderating influences as this is very difficult to prove empirically. Also note the inclusion of factors that affect the dissatisfied consumer's perception of justice. It is argued that factors other than procedural, interactional and distributive justice impact on whether or not a dissatisfied consumer takes a complaint further. Other factors which must be taken into account are the speed of redress, the degree of redress sought, the type of redress sought (e.g. monetary or an apology), and the rank of the employee dealing with the voiced complaint. Note that the model includes "buying behaviour" and feedback loops (as indicated by the dotted lines) which are to signify that consumers do not buy in a vacuum: both past complaining and purchase experiences affect future buying behaviour. It may well be the case that, as suggested by an anonymous reviewer of this paper, a single dissatisfying experience with a product or service may be dismissed as an isolated incident, but a repetition over time (i.e. a cumulative experience) might well trigger a complaining behaviour. Therefore, a repetition of the problem may result in a more intense affective response. Clearly the temporal dimension of CCB is an area

ripe for future conceptual and empirical research. It is also shown on the model, through dotted lines that past levels of perceived justice will have an impact on future company/consumer, and marketplace/consumer, relationships. The four stages of the model are marked by vertical dashed lines.

CONCLUSION

Through a review of the current literature on complaining behaviour, this paper offers a comprehensive taxonomy and model of consumer dissatisfaction responses. The next step will involve the empirical testing of the model in order to assess its validity. As East (1998b) makes clear, we are still some way off from assessing the relative weight of each trigger of consumer dissatisfaction responses. Once we have established that, we can then assess how important are the triggers' various dimensions. Singh's study (1990) of a selection of the triggers of CDRs explained 55% of variance of consumers' complaining behaviour. It is hoped that by applying the taxonomy and model presented here, future research may be able to explain a great deal more of the variance in consumer complaining behaviour than has previously been reported in the literature.

REFERENCES

- Ajzen, I. (1991), "The Theory of Planned Behaviour," *Organisational Behaviour and Human Decision Processes*, 50, 179-211.
- Andreasen, A. "Consumer Complaints and Redress: What We Know and What We Don't Know," *The Frontier of Research in the Consumer Interest*, E. S. Maynes (ed.), Chapter 28, Missouri: American Council on Consumer Issues, 675-722.
- Andreasen, A. and A. Best (1977), "Consumers Complain - Does Business Respond?" *Harvard Business Review*, 55, 93-101.
- Bearden, W. O. and J. E. Teel (1983), "Selected Determinants of Consumer Satisfaction and Complaint Reports," *Journal of Marketing Research*, 20, (February), 21-28.
- Blodgett, J. G. (1994), "The Effects of Perceived Justice on Complainants' Repatronage Intentions and Negative Word-Of-Mouth Behaviour," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 7, 1-14.
- Blodgett, J. G. and D. H. Granbois (1992), "Toward an Integrated Conceptual Model of Consumer Complaining Behaviour," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 5, 93-103.
- Blodgett, J. G. and S. S. Tax (1993), "The Effects of Distributive and Interactional Justice on Complainants' Repatronage Intentions and Negative Word-Of-Mouth Intentions," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 6, 100-110.
- Cadotte, E. R., R. B. Woodruff and R. L. Jenkins (1987), "Expectations and Norms in Models of Consumer Satisfaction," *Journal of Marketing Research*, 24, (August), 305-314.
- Churchill, G. A. and C. Surprenant (1982), "An Investigation into the Determinants of Customer Satisfaction," *Journal of Marketing Research*, 19, (November), 491-504.
- Day, R. L. (1980), "Research Perspectives on Consumer Complaining Behaviour," *Theoretical Developments in Marketing*, C. Lamb and P. Dunne (eds.), Chicago: American Marketing Association, 211-215.
- Day, R. L. (1984), "Modelling Choices among Alternative Responses to Dissatisfaction," *Advances in Consumer Research*, T. C. Kinnear (ed.), 11, Ann Arbor, MI: Association for Consumer Research, 496-499.
- East, R. (1998a), "Identifying the Drivers of Complaining," *Proceedings of the 27th European Marketing Academy Conference*, P. Andressen (ed.), 399-418.
- East, R. (1998b), "The Bases of Consumer Complaining" In *Proceedings of the 1998 Annual Conference of the Academy of Marketing*, 178-181.
- Erevelles, S. and C. Leavitt (1992), "A Comparison of Current Models of CS/D," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 5, 104-114.
- Fails, S. L. and S. K. Francis (1996), "Effect of Demographic Variables on Self Concept of Positive and Negative Word-of-Mouth," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 9, 200-206.
- Feinberg, R. A., R. Widdows and R. Steidle (1996), "Customer (Dis)satisfaction and Delays: The Robust Negative Effect of Service Delays," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 9, 81-85.
- Folkes, V. S. and B. Kotsos (1986), "Buyers' and Sellers' Explanations for Product Failure: Who Done It?" *Journal of Marketing*, 50, (April), 74-80.
- Fornell, C. and N. Didow (1980), "Economic Constraints on Consumer Complaining Behaviour," *Advances in Consumer Research*, J. Olsen (ed.), 8, Ann Arbor, MI: Association for Consumer Research, 318-323.
- Foxman, E. R., P. V. Raven, and D. E. Stem (1990), "Locus of Control, Fatalism, and Responses to Dissatisfaction: A Pilot Study," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 3, 1-14.

- Behaviour*, 3, 21-28.
- Godwin, B. F., P. G. Patterson and L. W. Johnson (1995), "Emotion, Coping and Complaining Propensity Following a Dissatisfactory Service Encounter," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 8, 155-163.
- Goodwin, C. and I. Ross (1990), "Consumer Evaluations of Responses to Complaints: What's Fair and Why," *Journal of Consumer Marketing*, 7, 2, (Spring), 39-47.
- Hirschman, A. O. (1970), *Exit, Voice, and Loyalty. Responses to Decline in Firms, Organisations and States*, Cambridge, MA: Harvard University Press.
- Huefner, J. C. and H. K. Hunt (1994), "Extending the Hirschman Model: When Voice and Exit Don't Tell The Whole Story," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 7, 267-270.
- Hunt, H. D. and H. K. Hunt (1990), "Customer Grudgeholding: Further Conceptualisation and Analysis," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 3, 117-122.
- Kolodinsky, J. (1993), "Complaints, Redress, and Subsequent Purchases of Medical Services by Dissatisfied Consumers," *Journal of Consumer Policy*, 16, 193-214.
- Kolodinsky, J. and J. Aleong (1990), "An Integrated Model of Consumer Complaint Action Applied to Services: a Pilot Study," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 5, 36-44.
- Kowalski, R. M. (1996), "Complaints and Complaining: Function, Antecedents and Consequences," *Psychological Bulletin*, 119, 2, 179-196.
- Lapidus, R. S. and L. Pinkerton (1995), "Customer Complaint Situations: An Equity Perspective," *Psychology and Marketing*, 12, (March), 105-122.
- Liefield, J. (1980), "Urban/Rural Consumer Expectations and Evaluations of Their Consumer Realities," *Refining Concepts and Measures of Consumer Satisfaction and Complaining Behaviour*, R. Day and H. K. Hunt (eds.), Bloomington, IN: Indiana University School of Business, 95-101.
- Malafi, T. N., M. A. Cini, S. L. Taub and J. Bertolami (1993), "Social Influence and the Decision to Complain: Investigations on the Role of Advice," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 6, 81-89.
- Miller, J. A. (1977), "Studying Satisfaction, Modifying Models, Eliciting Expectations, Posing Problems and Making Meaningful Measurements," *Conceptualisation and Measurement of Consumer Satisfaction and Dissatisfaction*, H. K. Hunt (ed.), Cambridge, MA: Marketing Science Institute, 72-91.
- Nantel, J. (1985), "Can the Self-monitoring Construct Improve Predictions of Consumer Complaining/Noncomplaining Behaviours?," *Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, H. K. Hunt and R. L. Day (eds.), Indiana University: Marketing Department.
- Office of Fair Trading (1986), *Consumer Dissatisfaction: A Report on Surveys Undertaken for the Office of Fair Trading*, London: Office of Fair Trading.
- Parker, R. G., R. Funkhouser and A. Chatterjee (1993), "Some Consumption Orientations Associated with Propensity to Complain," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 6, 113-117.
- Prakash, V. (1991), "Intensity of Dissatisfaction and Consumer Complaint Behaviours," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 4, 110-122.
- Raven, P. V. and E. R. Foxman (1994), "Responses to Dissatisfaction: A Cross-Cultural Comparison," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 7, 236-245.
- Richins, M. L. (1983), "An Analysis of Consumer Interaction Styles in the Marketplace," *Journal of Consumer Research*, 10, (June), 73-82.
- Rogers, J. C., S. Ross and T. G. Williams (1992), "Personal Values and Purchase Dissatisfaction Response," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 5, 81-92.
- Singh, J. (1988), "Consumer Complaint Intentions and Behaviour: Definitional and Taxonomical Issues," *Journal of Marketing*, 52, (January), 93-107.
- Singh, J. (1990), "A Typology of Consumer Dissatisfaction Response Styles," *Journal of Retailing*, 66, (Spring), 57-99.
- Singh, J. and R. E. Wilkes (1991), "A Theoretical Framework for Modelling Dissatisfaction," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 4, 1-12.
- Singh, J. and R. E. Wilkes (1996), "When Consumers Complain: A Path Analysis of the Key Triggers of Consumer Complaint Response Estimates," *Journal of the Academy of Marketing Science*, 24, 4, 350-365.
- Slama, M. and K. Celuch (1994), "Assertion and Attention to Social Comparison Information as Influences on Consumer Complaint Intentions," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 7, 1994.
- Slama, M., M. D'Onofrio and K. Celuch (1993), "Consumer Complaint Behaviors of Market Mavens," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 6, 175-180.
- Standop, D. (1991), "Manufacturers' Concern for Product Risks After Sale: Do Product Recalls Really Cause Consumer Dissatisfaction?" *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 4, 80-83.
- Technical Assistance Research Program (TARP) (1979), *Consumer Complaint Handling in America: Summary*

- of Findings and Recommendations*, Washington, D.C.: TARP.
- Tesser, A. and S. Rosen (1975), "The Reluctance to Transmit Bad News," *Advances in Experimental Social Psychology*, 8, 193-232.
- Tse, D. K. (1990), "Attributing Performance Discrepancy: Dimensions, Causes and Effects," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 3, 105-111.
- Tse, D. K. and P. C. Wilton (1988), "Models of Consumer Satisfaction Formation: An Extension," *Journal of Marketing Research*, 25, (May), 204-212.
- Woodruff, R. B., E. R. Cadotte and R. L. Jenkins (1983), "Modelling Consumer Satisfaction Processes Using Experienced-Based Norms," *Journal of Marketing Research*, 20, (August), 296-304.
- Woodruff R. B., D. S. Clemons, D. W. Schumann, S. F. Gardial and M. J. Burns (1991), "The Standards Issue in CS/D Research: A Historical Perspective," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour*, 4, 103-109.
- Yi, Y. (1990) "A Critical Review of Consumer Satisfaction," *Review of Marketing*, American Marketing Association: Chicago, 68-123.

Send correspondence regarding this article to:

Jonathan Boote
Department of Marketing and International Business
Luton Business School
University of Luton
Park Square
Luton, Bedfordshire LU1 3JU
UNITED KINGDOM

AN EXPLORATORY STUDY OF INTERNATIONAL STUDENT SATISFACTION

Mary L. Carsky, The University of Hartford
Susan Coleman, The University of Hartford

ABSTRACT

The specific objectives of this study, conducted to measure international student satisfaction with their university experience, were to assess the effectiveness of services provided by the campus Office of International Student Services, to identify international students' problems in adjusting to campus life on this campus, and to measure international students' satisfaction with the university environment. The major finding was that students' perceptions of pre-arrival information were more important determinants of satisfaction than were some post-arrival experiences. Neither the nature and severity of adjustment problems nor the level of participation in campus life were found to be significant predictors of satisfaction.

INTRODUCTION

Over the past ten years, medium-sized private and public universities have engaged in more aggressive recruitment of international students. International students enrich the university environment. Their presence internationalizes the campus in a way that could not be achieved simply by adding cross cultural-courses. American students have first-hand opportunities to become acquainted with people from around the world, to understand their cultures, and to prepare themselves for operating in the global economy. In addition, the declining college-age population in the United States has left empty seats in the classroom that international students fill.

Building a culturally diverse student population takes time. In order to attract students from countries around the world, it is essential to develop a critical mass from a country. It is necessary to first attract approximately ten students, and from this base, more will come. Of course, the base will expand only if the students are satisfied with both academic and non-academic opportunities and services of the university.

This study was conducted to generate information to guide the Office of International

Student Services in serving International Students. The office is responsible for sending *pre-arrival* information, conducting international student orientations, and handling problems on a case-by-case basis. Thus, the study was designed to examine the effectiveness of the services of the Office of International Student Services.

Four focus groups were conducted to identify the most salient issues and concerns of international students and to determine whether problems varied among students of different nationalities. Each focus group was comprised of students from a different country (Korea, Italy, India and Turkey). Both men and women participated in each of the groups. Cultural differences were identified. Italian, Indian and Turkish students felt the pre-arrival information was insufficient and inaccurate. Korean and Indian students indicated problems with academic advising, and Koreans experienced language difficulties.

Italian students said they wanted to live on campus to learn the American culture. The remaining groups indicated that campus housing was too expensive, there was not enough privacy, or that they wanted to live with students from their own culture. University organized activities were not favored by any of the groups, but for different reasons. For example, Italians wanted to explore social activities on their own; they felt uncomfortable at university organized activities. Other groups tended to say that not enough activities were directed towards them.

This study measured international students' satisfaction with the educational environment. It was an exploratory study that differed from tests of the CS/D models in several ways:

1. The study examined the extent to which multiple components of the educational experience influenced international student satisfaction. The study did not follow the traditional CS/D process models (Erevelles and Leavitt 1992). Expectation, a key construct of the CS/D models, was not measured.
-

2. The study measured satisfaction with an on-going complex service. By contrast, many tests of CS/D examined satisfaction with a product purchase and/or use experience, generally considering a single transaction or consumption experience (see Perkins 1992).

3. Student satisfaction with the educational experience may be more akin to measuring service quality. The service quality literature (e.g. Zeithaml, Parasuraman and Berry 1990, 1993; Cronin and Taylor 1992, 1994) suggests that evaluation over many transactions is a measure of service quality rather than consumer satisfaction.

The study was conceptualized, in part, in accordance with Spreng, MacKenzie and Olshavsky (1996). These authors contended that satisfaction with information about the product or service as well as satisfaction with product/service attributes determine overall satisfaction. Thus, we measured perceptions of the information sent to students prior to arriving in the United States.

RESEARCH OBJECTIVES

This study was conducted to measure international student satisfaction. Specific objectives of the study were to assess the effectiveness of services provided by the Office of International Student Services, to identify problems in adjustment to campus life in this country, and to measure international students' satisfaction with the university environment.

If the *pre-arrival* information was accurate and complete, one might expect the students to have fewer problems adjusting to the university environment. Further, those who experienced fewer problems in adjustment would be more satisfied. Likewise it would be expected that those who participated in campus activities thus becoming part of the community would be more satisfied. Therefore it was hypothesized that:

H1: Adjustment to the university would be a more important determinant of satisfaction than the pre-arrival information.

H2: Students who had fewer problems

adjusting to the university environment would be more satisfied.

H3: Students who participated in campus activities would be more satisfied.

It would be expected that perceived quality of university services would moderate international student satisfaction. Both the academic and non-academic environments would be important to students. Course work and interaction with faculty would be salient. Therefore, it was hypothesized that

H4: Students who perceived the quality of academic services to be high would be more satisfied.

One of the key issues that surfaced during the focus groups was housing. Students expressed concerns about on-campus housing. Many chose to live off-campus where they believed the cost to be lower and the quality-of-life better. Therefore, it was hypothesized that:

H5: Students who perceived the quality of housing on campus to be high would be more satisfied with the university.

METHODOLOGY

A four page booklet style questionnaire was developed. The questionnaire was divided into six sections, each focusing on a different aspect of international student adjustment and satisfaction. The questionnaire was pretested on twenty-six students in a marketing research class which included two internationals. Revised questionnaires were mailed to 500 of the international students at the university. A low response rate prompted an intercept administration. International students were contacted in the residence halls, academic buildings and the student union. An effort was made to survey a representative sample of students from the various countries.

MEASURES

Multiple indicators were used to measure

student perceptions of the pre-arrival information, adjustment to the university, perceived performance of the university and satisfaction with the educational experience. Scales were constructed to measure students' adjustment. These included problems upon arrival and the level of participation in campus activities.

Perceptions of the Pre-Arrival Information

Respondents were asked to rate the pre-arrival information on topics including tuition and fees, transportation, housing, banking, health insurance, and academic issues such as program offerings and registration processes. Ratings were obtained on five-point scales (poor-excellent). The ratings on the seven items were summed.

Measures of Adjustment

Adjustment to campus life was measured by a series of questions on the nature and severity of problems encountered upon arrival and on their participation in campus activities. Students were asked to rate the difficulties of finding housing, banking services, transportation, and insurance. They were also asked about difficulties with communication and cultural adjustment.

A second measure of *adjustment* was the level of participation in campus activities. Information on participation was solicited through a series of questions on awareness of and participation in social activities and attendance at the International Student Orientation.

Measures of Performance

Perceived performance was measured by students' interaction with faculty and perceptions of housing. Respondents were asked who assisted them in resolving academic issues such as course selection and registration. Particular emphasis was placed on the quality of students' relationships with their academic advisors.

Because housing issues were among the most important to students in the focus groups, several questions addressed housing issues. Respondents were asked their impressions of various attributes of on-campus housing. They were also asked whether they chose to live on or off campus, and

the primary reason for choosing to live off campus.

Measures of Satisfaction

Three questions were used to assess global satisfaction with the university. These three were derived from the literature on measuring consumer/customer satisfaction (e.g., Hunt 1989; Hausknecht 1991). Students were first asked how satisfied they were with their educational experience, a global measure of satisfaction. Next they were asked whether they would choose the university again, a measure of intention which indicates that the previous purchase was satisfactory (Hunt 1989). The final question asked the likelihood of recommending the university to others in their home country. These three were summed to create an interval-like scale.

RESULTS

Completed, usable surveys were received from 147 students, representing a 29 percent response rate. Respondents represented international students from 22 different countries in Asia, Africa, Europe and Central/South America. Representation from the various regions was relatively proportional. Approximately two-thirds of the respondents were graduate students enrolled in the Business School. Half were completing their first year at the university. The majority (67.2%) lived off-campus.

Reliability of the Scales

Cronbach's alpha was used to assess the reliability of the scales developed to measure the constructs. Alpha for all scales was above .70 as shown in Table 1. The table shows the number of items used to construct each scale along with alpha.

International Student Satisfaction

Satisfaction was measured by three questions. First, students were asked how satisfied they were with their educational experience. Second, they were asked whether they would choose the university again; and third, they were asked

whether they would recommend the university to others in their home countries. As shown in Table 2, the intercorrelation of the three measures of satisfaction (SAT, AGAIN and RECOM) as well as their individual correlations with the composite measure of satisfaction (TOTSAT) were positive and statistically significant.

Table 1
Cronbach's Alpha for All Scale

Scale to measure:	n of items	alpha
Pre-arrival information (satisfaction with information)	6	.778
Problems upon arrival (Adjustment)	10	.729
Perception of housing (Performance)	6	.777
Helpfulness of academic advisor (Performance)	4	.922
Satisfaction	3	.761

Table 2 presents the correlations among the independent variables and between the independent variables and the measures of satisfaction. The perceived quality of the academic assistance, whether or not they liked the campus housing and the pre-arrival information are shown to be significantly correlated with the composite measure of satisfaction (TOTSAT) and with all three components of the satisfaction measure.

The number of campus activities in which respondents participated (PARTIC) is significantly

related only to whether the respondent would recommend the university to others; however, the relationship is negative. The number and severity of the problems encountered (PROBS) is related to whether the respondent would choose this university again. The length of time the respondent has been at the university is unrelated to satisfaction; it is only related to the level of participation in campus activities. The longer the student had been at the university, the more campus activities s/he has attended.

Demographic characteristics of the respondents were not related to any of the satisfaction measures nor were they related to the independent variables. Neither gender nor marital status nor home country were related to any of the measures.

Hypotheses Tests

The first three hypotheses were concerned with students' adjustment to the university. It was hypothesized that problems encountered during adjustment and participation in campus activities would be more important than the *pre-arrival* information determining satisfaction. These hypotheses were tested simultaneously using GLM regression on SAS. The three independent variables of pre-arrival information, problems encountered, and participation in campus activities were entered. The composite measure of satisfaction (TOTSAT) was the dependent variable. The equation was significant at $p = .003$ ($F=6.81$). The equation explained a very small

Table 2
Correlations Between Variables to Measure Satisfaction with the University Environment

	AGAIN	RECOM	TOTSAT	PARTIC	ARIVINF	PROBS	ACADSE	ATUNIV	HOUSING
How SATisfied	.4927***	.5408***	.7459***	.0433	.3377***	.0996	.2542**	.0712	.3386***
Would come AGAIN		.6240***	.8125***	.0983	.2105*	.2248**	.3179***	.0141	.2462**
Would RECOM			.9066***	.1809	.3557***	.1632	.3307***	.1121	.2045*
TOTSATisfaction measures				.1170	.3634***	.2043*	.3633***	.0483	.3219***
#of activities PARTICIPated in					.0655	.1130	.0287	.2322**	.0354
Perception of (pre)ARRIVAL INFO.						.1581	.2821**	.0087	.2687**
PROBLEms encountered							.1550	.0232	.0169
ACADSER (academic advisor helpfulness)								.0547	.2033
How long ATUNIVersity									.0702
Perception of HOUSING									

*p < .05 **p < .01 ***p = < .001

proportion of the variation in satisfaction ($R^2=.136$). The parameter estimates and significance of the independent variables are given in Table 3.

Table 3
Regression Coefficients for Adjustment

Predictor	coef.	t	p
Intercept	7.457	8.91	.0001
Pre-arrival info.	0.120	3.82	.0002
Problems on arrival	-0.035	-1.74	.0946
Participation	-0.076	-0.61	.5434

As shown in the table, only pre-arrival information ($p = .0002$), was significant. Thus, the three hypotheses were not supported. Contrary to the hypotheses, perceptions of the pre-arrival information were more influential on satisfaction than the number and severity of the problems encountered.

The fourth and fifth hypotheses posited a relationship between students' perceptions of selected academic and non-academic components of their educational experience and satisfaction. The quality of academic advising and of on-campus housing were tested simultaneously using the GLM procedure. Consistent with the model posited by Spreng, MacKenzie, and Olshavsky (1996), satisfaction with the *pre-arrival* information was an important predictor of satisfaction. The equation was statistically significant ($F=11.95$; $p=.0001$). The three variables explained nearly one fourth of the variance in satisfaction ($R^2 = .2201$). The parameter estimates and significance of the independent variables are given in Table 4.

Table 4
Regression Coefficients for Performance

Predictor	coef.	t	p
Intercept	5.104	10.68	.0001
Pre-arrival info.	0.080	2.53	.0130
Academic Service	0.086	3.15	.0021
Housing	0.055	2.54	.0120

As shown in the table, all three independent variables were significant. The perceived quality of housing was an important determinant of

satisfaction. In order to determine whether the perception of housing was mediated by students choice of on-campus or off-campus housing, a GLM model including the dummy variable for choice of housing was run. With this addition, the housing variable ceased to be significant and the R^2 dropped to .2014.

The independent variable of pre-arrival information was found to be an important moderator of satisfaction with the selected measures of performance. When pre-arrival information was removed from the equation, the proportion of explained variance dropped substantially ($R^2=.1858$). Thus, perceptions of the accuracy and completeness of the pre-arrival information, moderated student satisfaction along the dimensions examined.

DISCUSSION

The primary purpose of this study was to measure international student satisfaction. The questions asked and the resultant measures used for these analyses were able to explain less than one-fourth of the variance in satisfaction. Many variables that would be expected to explain differences failed to do so. Several plausible explanations for the results are offered.

1. The data for this study was gathered from 147 international students at a mid-sized university in the northeastern United States. The students, from the five continents and countries such as Finland, France, Argentina, Romania, Tanzania, Thailand, and Uganda, would be expected to have vastly different attitudes, values, and perceptions.. No more than 20 students came from any single country. Thus, the sample of 147 students might have been too heterogeneous.

2. Some of the questions may have been misunderstood or interpreted differently by the respondents. While all students are able to read and write in the English language, terms used in the survey might have had different meanings to some students.

3. The respondents were currently enrolled at the university and still involved with the services of the institution. Their attitudes and perceptions might differ from one day to the next, depending on that day's experiences.

4. Students might have given responses that

were socially desirable. In some cultures, it is considered impolite to express dissatisfaction. Conversely, some students might not want to reveal a problem. For example, when asked whether they had difficulties with the language, nearly half (47%) said that this was not a problem. Less than 15 percent indicated that language was a serious problem.

5. The study was designed to gather information on factors that could be somewhat controlled by the Office of International Student Services. These might not have included some of the most salient student issues.

6. The composite measure of satisfaction used in this study was developed by combining a global measure of satisfaction, a measure of intentions, and a third measure, the willingness to recommend the university to others. Although the three were intercorrelated at .5 or above, they clearly tap different dimensions of satisfaction. Students who reported that they would come to this university again, might not have been fully satisfied. Conversely, those who were satisfied, might be hesitant to recommend the school to others from their country.

The Findings on International Student Satisfaction

The major finding of this study was that students' perceptions of pre-arrival information, were more important determinants of satisfaction than some post-arrival experiences. Neither the nature and severity of adjustment problems nor the level of participation in campus life were found to be significant predictors of satisfaction.

Recruiters of international students recognize the importance of obtaining a "critical mass" of ten or more students in order to attract more. International students, though they come to this country for an American education, prefer to interact with fellow countrymen with whom they share a common language and a common value perspective. It would be expected that international students who have more countrymen with whom to socialize would be more satisfied. This was not the case. We found no differences in satisfaction between students whose countries were well represented and those who had three or fewer classmates from their home countries.

Alternative explanations may be offered for the failure to find an effect of problems encountered or level of participation on student satisfaction. Several international students said that when moving to a new country, particularly a country whose culture is vastly different, adjustment problems are anticipated. Students reported relying heavily on friends and family from their own country for assistance with housing, transportation, banking, medical care and other life essentials. They also reported relying primarily on their friends for advice on course selection and other academic matters.

After considering international students' patterns of behavior, particularly their preference for socializing with fellow countrymen, it is no surprise that the level of participation in campus activities was not a determinant of satisfaction. While it might have been expected that those students who "fit in" with the campus lifestyles, would be more satisfied, international students tastes and preferences for social activities are often quite different from that of American students--for whom the activities are designed. Many internationals are not interested in the types of social activities targeted to Americans. Yet, among those who indicated dissatisfaction, many failed to participate in activities planned in accordance with student requests and organized by the Office of International Student Services.

The findings of this study support Spreng, MacKenzie and Olshavsky (1996) in their contention that satisfaction with the "product" and satisfaction with the "information about the product" are distinct constructs. They stated that information satisfaction is more important for products with "lots of experience attributes" because the consumer is dependent upon someone else for information (Spreng, MacKenzie and Olshavsky 1996:28). Certainly an education, particularly in a foreign environment, is replete with experience attributes.

EPILOG

Based on the results of this study, the *pre-arrival* information sent to international students was revised to include more complete information about life in the northeastern United States, housing, transportation, insurance, and other life

issues. The following year, the Office of International Student Services found a reduction in the number and severity of adjustment problems reported by first year students.

REFERENCES

- Cronin, J. Joseph and Steven A. Taylor (1992), "Measuring Service Quality: A Reexamination and Extension," *Journal of Marketing*, 56, 55-68.
- Cronin, J. Joseph and Steven A. Taylor (1994), "SERVPERF Versus SERVQUAL: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality," *Journal of Marketing*, 58 (January), 111-124.
- Erevelles, Sunil and Clark Leavitt (1992), "A Comparison of Current Models of Consumer Satisfaction/Dissatisfaction," *Journal of Consumer Satisfaction/Dissatisfaction & Complaining Behavior*, 5, 104-114.
- Hausknecht, Douglas R. (1990), "Measurement Scales in Consumer Satisfaction/Dissatisfaction." *Journal of Consumer Satisfaction/Dissatisfaction & Complaining Behavior*, 3, 1-11.
- Hunt, H. Keith (1989), "Customer Satisfaction, Dissatisfaction and the Consumer Interest," In *The Frontier of Research in the Consumer Interest*, E. Scott Maynes, ed., 731-748.
- Perkins, Debra (1992), "A Consumer Satisfaction, Dissatisfaction and Complaining Behavior Bibliography: 1982-1990," *Journal of Consumer Satisfaction/Dissatisfaction & Complaining Behavior*, 4, 194-227.
- Spreng, Richard A., Scott B. MacKenzie and Richard W. Olshavsky (1996), "A Reexamination of the Determinants of Consumer Satisfaction," *Journal of Marketing*, 60 (July), 15-32.
- Zeithaml, Valarie A., Leonard L. Berry and A. Parasuraman (1990), *Delivering Quality Service: Balancing Customer Perceptions and Expectations*. New York: The Free Press.
- Zeithaml, Valarie A., Leonard L. Berry and A. Parasuraman (1993), "The Nature and Determinants of Consumer Expectations of Service," *Journal of the Academy of Marketing Science*, 21 (Winter), 1-12.

Send correspondence regarding this article to:

Mary L. Carsky
The University of Hartford
Barney School of Business and Public Administration
West Hartford, CT 06117 USA

TOLL-FREE CUSTOMER SERVICE TELEPHONE LINES: ARE THEY ACCESSIBLE FOR SPANISH-SPEAKING CONSUMERS?

Dennis E. Garrett, Marquette University
Carlos Torelli, Mavesa, S.A.
Amy Olson Gunderson, Marquette University

ABSTRACT

A survey of companies whose products list toll-free customer service numbers shows that a majority of these companies' service centers have Spanish-speaking representatives who can assist consumers who prefer to speak Spanish rather than English. However, additional analysis reveals several formidable obstacles in the customer service process that discourage Spanish-speaking consumers from utilizing these services. Because the number of Spanish-speaking consumers in the United States is growing rapidly, consumer affairs managers should endeavor to improve their companies' toll-free customer service systems. Based on the findings from this study, it is recommended that consumer affairs managers implement and promote separate toll-free customer service numbers specifically for their Spanish-speaking customers.

INTRODUCTION

During the past decade there has been a dramatic transformation in the manner in which dissatisfied consumers communicate their complaints to companies whose products they have purchased. With the advent of toll-free telephone systems, many progressive companies, particularly in the consumer products arena, have encouraged customers to call their service representatives directly, rather than taking the time and effort to write a letter, to express their complaints or questions (Garrett and Meyers 1996; Martin and Smart 1994). By all accounts, it appears that these toll-free customer service telephone numbers have proven to be beneficial for both consumers and consumer affairs managers as they endeavor to maximize customer satisfaction (SOCAP 1992).

While these toll-free customer service telephone systems appear to work very well for most consumers, concerns remain regarding the status of vulnerable consumers. Andreasen and Manning (1990, p. 13) define vulnerable consumers "as those who are at a disadvantage in

exchange relationships where that disadvantage is attributable to characteristics that are largely not controllable by them at the time of the transaction." According to Andreasen and Manning (1990), examples of vulnerable consumers include children, the elderly, the uneducated, the structurally poor, the physically handicapped, ethnic and racial minorities and those with language problems. Additionally, they note that despite the fact that vulnerable consumers have much to lose when they are dissatisfied with their purchases, vulnerable consumers appear to be less likely than other consumers to pursue constructive complaining options. Although little research has specifically analyzed vulnerable consumers' complaining behavior, Andreasen and Manning (1990) suggest that vulnerable consumers may not complain because they have fewer options for complaining.

OBJECTIVE

The purpose of this research project is to investigate how well toll-free customer service telephone systems function for one specific group of vulnerable consumers, Hispanic consumers in the United States who speak Spanish rather than English. Andreasen and Manning (1990) specifically highlighted Hispanics as one of the notable vulnerable groups that has been largely ignored in the consumer satisfaction research field. Therefore, this study will analyze whether or not toll-free customer service telephone systems, which apparently serve quite well the needs of English-speaking consumers, can accommodate the Spanish language needs of Hispanic consumers who call with questions or complaints about their purchases.

We will begin by discussing in more depth the importance of researching Spanish-speaking Hispanic consumers. Then the specific research questions that are investigated in this study will be presented, followed by the methods used to answer these questions. Finally, the results of this study will be presented and discussed.

THE IMPORTANCE OF RESEARCHING SPANISH-SPEAKING HISPANIC CONSUMERS

Three factors underscore the importance of investigating how well toll-free customer service telephone systems function for Spanish-speaking Hispanic consumers. We will discuss first the growing size and economic significance of the Hispanic market in the United States. We will then focus on the language preferences of Hispanic consumers. Finally, the lack of research devoted to this topic in the existing body of consumer satisfaction literature will be highlighted.

Size and Economic Significance of the Hispanic Market in the United States

The first reason why this research issue is important deals with the rapidly growing size and economic value of the Hispanic market. Due to their higher birth rates and immigration patterns, the number of Hispanics in the United States has increased dramatically. The Hispanic population grew from about 22.5 million in 1990 to about 27 million in 1995, which was about 10% of the total U.S. population (Hamstra 1996). If current projections are correct, there will be 31 million Hispanics in the U.S. by 2000 and 41.4 million by 2010, at which time Hispanics will outnumber African-Americans as the largest minority group in the United States (Hamstra 1996; Hitt 1998). Also, the number of Hispanic households is predicted to grow from 4.1 million in 1983 to over 9 million by 2000, more than double the expected increase for African-American households (Valdes and Seoane 1995).

Correspondingly, spending by Hispanics is playing an increasingly significant role in the United States economy. In terms of consumer expenditures, spending by Hispanics in the U.S. market is estimated to be \$348 billion in 1997, which represents a 66% increase from 1990 levels (Wynter 1997a). In comparative terms, African-American expenditures increased by 54% from 1990 to an estimated \$469 billion in 1997 (Wynter 1997a). As might be expected, many companies have recognized this growth in the Hispanic market and have attempted to develop marketing campaigns to appeal to this potentially lucrative

segment (Millman 1998; Petrozzello 1995; Wynter 1997b; Zbar 1995). According to *Hispanic Business* magazine, advertising spending in the Hispanic market exceeded \$1 billion in 1995 (Hamstra 1996). Therefore, because of the growing size and economic clout of Hispanic consumers in the United States, it is imperative to research how well companies' toll-free customer service lines function for them.

Language Preferences of Hispanics in the United States

A second reason why this study is potentially important relates to the language usage patterns and preferences of Hispanics in the United States. Valdes and Seoane (1995, pp. 12-13) assert:

Of all social traits, language is perhaps the most distinguishable characteristic of any culture, and probably the last one any immigrant group will give up. As has been documented extensively, given the choice, most first-generational U.S. Hispanic adults speak in Spanish rather than English, even if they are bilingual. Spanish-speaking language usage is sure to continue well into the future if current immigration projections hold.

According to the 1990 census, of the 6.8 million Spanish-speaking households in the United States, 1.6 million were classified into the census definition of "linguistically isolated", or Spanish-dominant households (Valdes and Seoane 1995). The 1990 census also classified 4.5 million Spanish speakers age 5 and over as "linguistically isolated" because they spoke English "not well" or "not at all" (Valdes and Seoane 1995). Also, a study by Strategy Research, using 1990 census numbers and door-to-door interviews nationwide with 4,800 Hispanics, determined that 81% of Hispanics are most comfortable speaking Spanish, while only 17% prefer English and 2% use both languages (Fisher 1994).

Research has further shown that preference for the Spanish language exists even among highly acculturated Hispanics. Strategy Research found that among Hispanic adults 18 or older in the United States, 19% are highly acculturated, while 47% are partially acculturated and 34% are

relatively unacculturated (Fisher 1994). But even among highly acculturated Hispanic adults, Spanish was the language of greatest comfort and the language spoken most often at home. Of the highly acculturated Hispanic adults, 53% were most comfortable speaking Spanish, while 41% preferred English and 6% used both.

Moreover, it does not appear that this use and preference for Spanish will decline in the near future. According to DRI/McGraw-Hill, the number of Hispanics over age 5 who speak Spanish in the home will grow from 15.3 million in 1990 to 21 million in 2000 and 26.7 million by 2010 (Fisher 1994).

Because these data indicate that many Hispanic consumers in the United States are not comfortable speaking English, this further supports the need to determine if companies' toll-free customer service lines can accommodate the needs of Spanish-speaking Hispanic consumers. If these customer service centers are not able to meet the needs of Spanish-speaking consumers, these consumers are likely to experience significantly higher levels of consumer dissatisfaction. In turn, their increased level of dissatisfaction is likely to generate long-term financial problems for those companies that do not have customer service systems that can receive and respond to these consumers' complaints.

Research Regarding Hispanic Consumer Satisfaction

The third and final factor that highlights the need for this research is the relative paucity of prior attention to this topic. While the growing importance of the Hispanic market in the United States has generated a number of articles and research studies during the past two decades regarding Hispanic consumer behavior, virtually no prior research has concentrated on Hispanic consumers' use of toll-free customer service lines.

Early on in the development of Hispanic consumer behavior research, many practitioner-oriented articles sought to explain why marketers must address the unique needs and characteristics of Hispanic consumers (Astor 1981; Moya 1985; Segal and Sosa 1983; Valencia 1983). The general message of most of these articles was that marketers must recognize that Hispanic consumers

are indeed different from Anglo consumers along a number of dimensions, particularly their preference for promotional messages in Spanish rather than English. It appears that marketers who now attempt to reach the Hispanic market have accepted this message. In their survey of executives whose firms used advertising targeted at the Hispanic community, Albonetti and Dominguez (1989) found that the majority of these executives agreed that Spanish-language advertising is necessary for promoting to the Hispanic market in the United States.

The growing influence of the Hispanic market has also spawned a stream of academic research articles that have empirically investigated more thoroughly the nature of Hispanic consumer behavior. This research can be broadly divided into three major areas. First, some studies have concentrated on the shopping behavior of Hispanic consumers, particularly their degree of brand loyalty, as compared to non-Hispanic consumers (e.g., Mulhern and Williams 1994; Medina, Saegert, and Gresham 1996; Nicholls, Roslow, and Dublisch 1997; Saegert, Hoover, and Hilger 1985; Shim and Gehrt 1996; Wallendorf and Reilly 1983). Because Hispanics come from a variety of cultural backgrounds, a second group of studies has investigated potential differences in behavior among different categories of Hispanic consumers (e.g., Deshpande, Hoyer, and Donthu 1986; O'Guinn and Meyer 1983; Valencia 1989).

Finally, more pertinent to this study, a third set of research investigations has concentrated on Hispanic consumers' level of satisfaction with their purchases and their subsequent complaining behavior. In one of the earliest studies in this area, Villareal-Camacho (1983) reported that Mexican-Americans, as compared to Anglos, were less likely to complain and more likely to prefer exiting. In a more recent study, Diaz-Knauf, Schutz, and Almeida (1992) evaluated Hispanic consumers' satisfaction with a variety of categories of service providers. They found that Hispanic consumers were most satisfied with barbers/beauticians, health/fitness, dry cleaners, and appliance repair, and least satisfied with real estate, building contractors, and lawyers. Cornwell, Bligh, and Babakus (1991) investigated dissatisfied Mexican-American consumers' use of the Better Business Bureau. In contrast to national

complaint statistics for the Better Business Bureau, they discovered that far fewer complaints from Mexican-Americans focused on the area of unsatisfactory service. In an analysis of consumers' satisfaction with rainchecks as a solution to out-of-stock situations, Kelly, Huefner, and Hunt (1991) found that Hispanics, as compared to Whites, Blacks, and Orientals, were more satisfied with the use of rainchecks. Finally, Greene (1989) found that Hispanics who live in suburban communities are generally satisfied with their municipal services.

While these studies have added greatly to our understanding of Hispanic consumer behavior, the issue of Hispanic consumers' use of toll-free customer service numbers has received little research attention. The only prior research that attempted to address this issue was a national survey of toll-free customer service centers' activities and characteristics by the Society of Consumer Affairs Professionals (SOCAP 1992). This study reported that 32% of these centers have a bilingual representative. However, the significance of this finding is limited because SOCAP's survey included responses from companies in a broad variety of industries, many of which may not generate significant numbers of calls from Spanish-speaking consumers (e.g., "industrial products"). Also, SOCAP's research report did not specifically state if Spanish was the language that these "bilingual" representatives could speak. Most importantly, this study only asked if companies had bilingual representatives available. This study did not probe more deeply to consider how the entire toll-free customer service telephone system functions for consumers who do not speak English. For instance, if package instructions for calling a customer service number are written only in English, what is the likelihood that Spanish-speaking consumers will even be aware that this toll-free customer service number can be used to speak to a Spanish-speaking representative? Or, if Spanish-speaking consumers call a toll-free customer service number and are greeted by an automated service menu in English (i.e., "If you have a question about a product you bought, please press 2"), consumers may be unable to understand these instructions. Therefore, the question of how well toll-free customer service telephone systems serve the needs

of Spanish-speaking consumers remains unanswered.

In summary, the Hispanic market is a large, growing, and potentially lucrative market that more and more companies are endeavoring to target. However, the data clearly show that many Hispanic consumers are unable to speak and understand English. For this reason, it is vitally important to investigate how well the complaint management systems put in place in recent years by consumer products companies work for Hispanic consumers who are not fluent in English. If these consumers are not able to utilize the current customer service systems that most companies offer, this raises important questions that must be addressed. For instance, what potentially could companies do to modify their customer service telephone systems to provide better services for their Spanish-speaking customers? Also, if the current customer service telephone systems are not accessible for Spanish-speaking consumers, should governmental regulations be enacted to force companies to meet the needs of this segment of vulnerable consumers?

RESEARCH QUESTIONS

The fundamental objective of this study is to understand more thoroughly how accessible toll-free customer service centers are for those Hispanic consumers who are unable to speak English. This question of accessibility involves more than just determining if Spanish-speaking customer service representatives are available in service centers. Therefore, this study will investigate the following three-step process that Spanish-speaking consumers must undertake to express their complaints to customer service representatives.

Customer service information on product packaging. The first step in this process involves consumers becoming aware of the existence of toll-free customer service centers maintained by companies. Obviously, if consumers are not aware that companies provide toll-free customer service telephone lines for their use, they will not use them to communicate their complaints or questions. The following two questions address this issue of consumer awareness:

- 1) Do products list on their packaging a toll-free customer service number that customers may call if they have questions or complaints?
- 2) If products list a customer service phone number on their packaging, which language is used to communicate this information to consumers?

Service center processing of customers' calls. The second step in this process deals with how consumers' calls to toll-free customer service centers are processed and routed before consumers actually speak with service representatives. If Spanish-speaking consumers call these service centers and find that they are instructed in English to wait on hold or to select a certain service option from an automated service menu, they may become discouraged and abandon their attempts to contact a service representative. The next four questions focus on this issue:

- 3) When consumers call toll-free service centers, how often are their calls placed on hold?
- 4) If customers' calls to service centers are placed on hold, which language is used to instruct callers about this holding process?
- 5) When consumers call toll-free service centers, how often are they asked to use an automated service menu to route their call?
- 6) If customers receive an automated service menu when they call toll-free customer service centers, which language is used to explain to callers this routing and selection process?

Spanish-language services available at service centers. The third and final step in this process involves actual communication interactions that take place between consumers and service representatives. When Spanish-speaking consumers call toll-free customer service centers, it would be easier for them to communicate their complaints or questions if they are able to speak with service representatives who are fluent in Spanish. The final question in this study addresses this point:

- 7) What percent of toll-free customer service centers have Spanish-speaking service representatives on staff to meet the language needs of Spanish-speaking consumers who call?

METHODOLOGY

In this section, we will describe our choice of research setting, sampling procedures, and data collection process for this study.

Choice of Research Setting

We chose to focus on those products which Hispanic consumers typically purchase in grocery stores. While toll-free customer service numbers are now used by companies in a wide variety of industries (Garrett and Meyers 1996; Martin and Smart 1994; SOCAP 1992), products sold in grocery stores were specifically selected for analysis in this study for two primary reasons. First, a number of studies have shown that, compared to other groups of consumers, Hispanics spend a comparatively larger portion of their income on food items (Astor 1981; Fisher 1990; Segal and Sosa 1983; Valdes and Seoane 1995; Vargas 1997). According to a 1995 survey by the Bureau of Labor Statistics, Hispanics spend an average of \$3,370 annually on food for consumption at home, compared to \$2,803 for whites and \$2,367 for blacks (Vargas 1997). Because these products represent a relatively larger portion of Hispanics' budgets, a correspondingly larger proportion of Hispanics' questions and complaints are likely to focus on these same types of products.

Second, in addition to the traditional allotment of basic fresh and canned food products, most modern grocery stores now carry a broad assortment of other consumer products such as toiletries, pet supplies, laundry care products, and household cleaning products. Therefore, by selecting products from a modern full-line grocery store, a variety of the products that consumers use in their daily lives will be included in this study.

Sampling Procedures

Products were purchased from a full-line

grocery store in a large Midwestern city that is the home of over 40,000 Hispanics. This particular store was selected because it is located in a predominantly Hispanic section of this city as determined from local census data. The store, which is relatively new, attempts to cater to the local Hispanic population by posting most signs in both English and Spanish in the store. This store's commitment to the local Hispanic population is also evident because most of the service personnel are fluent in both English and Spanish. Previous research also shows that, contrary to the stereotype of Hispanics shopping mainly at local "Mom and Pop" grocery stores, Hispanics prefer to buy their groceries from large, full-line grocery stores such as the store selected for this study (Murphy 1997).

Two members of the research team randomly selected 145 products from the shelves of this store. All products in the store were eligible for selection except for those products in the frozen foods, fresh produce and fresh meat sections. These categories were excluded because the research team lacked facilities to properly store these products and prevent spoilage during the course of the research process. Also, these categories were excluded because many products in these sections typically do not have packaging that carries customer service information (e.g., bananas, packages of ground beef).

The 145 products that were selected came from the following product categories:

- 87: Food products
- 25: Personal care products (e.g., toiletries, cosmetics, medical care)
- 14: Household paper and plastic products (e.g., napkins, storage dishes)
- 10: Cleaning products (e.g., laundry soap, bathroom cleaners)
- 9: Pet care products (e.g., pet food, pet toys)

Data Collection Process

Each of the 145 products was first inspected to determine if a toll-free telephone number was listed on the product's packaging that customers could use to call the company if they had questions or complaints. If a toll-free telephone number was listed, the language used to convey this information on the package was noted. Then, if a

toll-free customer service number was listed, one member of the research team attempted to call this number. This researcher is a native of a South American country and is fluent in both Spanish, his native language, and English. When he placed these calls, the researcher recorded whether his calls were placed on hold and whether he was asked to use an automated service menu. He also noted the language that was used to explain to callers the holding process and automated service menus at these companies.

Up to three attempts were made to contact companies that listed a toll-free customer service number on their product packaging. When a customer service representative answered the call, the interviewer explained that the purpose of this call was to collect information regarding toll-free customer service numbers provided by companies for their customers. The interviewer explained that all responses obtained in this interview would be confidential and the company's name would not be included in subsequent reports. If the service representative agreed to participate, the interviewer first asked if the service center had Spanish-speaking customer service representatives available who could speak with Spanish-speaking consumers. The interviewer also asked the respondent how many calls the service center received per month from Spanish-speaking consumers regarding this particular product.

RESULTS

In this section the results obtained in this study will be reported in the three major areas being investigated: customer service information on product packaging, call processing of customer calls to toll-free service centers, and Spanish-language services available at toll-free customer service centers.

Customer Service Information on Product Packaging

The packaging of the 145 products selected for analysis in this study was examined to determine if consumers were provided with information regarding how to contact the companies if they had questions or complaints. This examination revealed:

- 71 products (49%) listed a toll-free customer service number.
- 74 products (51%) did not list a toll-free customer service number.

Next, the products were examined to determine the language that was used to instruct consumers on how to contact companies if they had questions or complaints. For those 71 products that provided information to consumers on how to call the company:

- 71 products (100%) provided information regarding toll-free customer service centers only in English.
- 0 products (0%) provided information regarding toll-free customer service centers in Spanish.

Call Processing of Customer Calls to Toll-Free Service Centers

When attempts were made to call the 71 companies that listed a toll-free customer service number on their packages, data were collected on the steps the caller must go through to make contact with a customer service representative. Of the 71 companies that were contacted:

- 60 companies (85%) did not place the caller on hold.
- 11 companies (15%) placed the caller on hold.

The researcher making the calls also noted the language that was used by these 11 companies to explain to consumers that their call was being placed on hold:

- 11 companies (100%) used English only for holding instructions.
- 0 companies (0%) offered Spanish language-holding instructions.

Automated service menus are an increasingly familiar mechanism by which companies attempt to more efficiently route consumers' calls to the most appropriate option (Grimm and Ward 1997). Of the 71 companies that were called:

- 41 companies (58%) presented the caller with an automated service menu.
- 30 companies (42%) did not present the caller with an automated service menu.

Once again, the researcher noted the language used by these 41 companies to explain to consumers how to use the automated service menu to route their calls to the most appropriate destination in the company:

- 41 companies (100%) used only English to explain their automated service menu to callers.
- 0 companies (0%) offered instructions for using their automated service menu in Spanish.

Spanish Language Services Available at Toll-Free Customer Service Centers

Next, attempts were made to determine if Spanish-language services were available at these 71 companies that provided toll-free telephone numbers for their consumers to call. To collect this information, it was necessary to speak directly with a member of the customer service center at each of these companies. Interviews were eventually completed at 62 of the 71 companies (87%). Interviews were not completed with the remaining nine companies (13%) because either the interviewer could not obtain an answer to his calls to the company or customer service personnel declined to participate in this study.

Of these 62 companies with whom interviews were completed:

- 7 companies (11%) reported that they did not have any ability to properly handle calls from Spanish-speaking consumers.
- 38 companies (61%) reported that they had at least one Spanish-speaking customer service representative on their "in-house" staff that could communicate directly with Spanish-speaking callers.
- 13 companies (21%) indicated that they utilized the services of an outside translation service to communicate with Spanish-speaking callers rather than having their own Spanish-speaking customer service representatives.

- 2 companies (3%) provided a system in which service representatives, once they recognized that callers preferred to speak Spanish, would play a recording in Spanish that directed consumers to leave a message so that a Spanish-speaking representative of the company could later call them back.
- 2 companies (3%) claimed they had the ability to properly handle calls from Spanish-speaking consumers but they declined to explain exactly the mechanism they used.

Before the results of this study are discussed, potential research limitations will be presented in the next section.

RESEARCH LIMITATIONS

There are a couple of limitations that must be weighed as the results of this study are interpreted. First, the focus of this study was limited to consumer products that are typically purchased from full-line grocery stores. Therefore, no conclusions can be drawn regarding the performance of customer service centers for other categories of consumer products, most notably consumer durables such as appliances, automobiles, and electronics products.

Second, products were selected from just one grocery store in one large Midwestern city. Therefore, the results from this study may not be generalizable to other grocery stores and other cities. However, given the national marketing and distribution patterns for many consumer products sold in grocery stores, the products selected for analysis in this study are probably quite similar to the products bought by Hispanics in other cities.

DISCUSSION

This section will first discuss how accessible customer service centers are for Spanish-speaking consumers. Next, discussion will focus on the implications of these findings for consumer affairs managers. Finally, the public policy issues related to the data found in this study will be addressed.

The Accessibility of Customer Service Centers for Spanish-Speaking Consumers

The overall objective of this study was to determine how well toll-free customer service centers developed by consumer products companies function for Spanish-speaking Hispanic consumers. The results found in this study present some contradictory evidence. On the one hand, the finding that 61% of those companies with toll-free customer service centers had Spanish-speaking service representatives on staff is promising. By itself, this suggests that perhaps the large number of Hispanic consumers who are not fluent in English may indeed be able to receive help from companies when they have questions or complaints regarding their purchases.

However, on the other hand, when the full process that Spanish-speaking consumers must undertake to receive service is evaluated, a far less positive picture emerges. As the results in this study clearly show, Spanish-speaking consumers must vault over several formidable barriers before they can even have the opportunity to speak with Spanish-language service representatives at these companies. First, because all of products in this study had directions on their packages regarding how to use their toll-free customer service numbers written only in English, it is doubtful if Spanish-speaking consumers would even understand that they should utilize these services when they have questions or problems. Second, even if they are aware that they can call these numbers to receive service, many Spanish-speaking consumers may be discouraged from actually continuing with this process when they are instructed in English to wait on hold or to use an automated service menu to route their call. The results in this study show that the use of holding instructions and automated service directions is a fairly common occurrence that could dissuade Spanish-speaking consumers from utilizing the services that may be available to them.

The seriousness of these obstacles for Spanish-speaking consumers is reinforced from additional information collected in this study. The interviewer also asked respondents to indicate the number or percent of calls from Spanish-speaking consumers received by their customer service centers. Respondents found it very difficult to

state a precise number or percentage, but they uniformly responded that they received "very few" calls from Spanish-speaking consumers. Given the obstacles Spanish-speaking consumers face, this low level of use of Spanish-language services in customer service centers is not surprising. Unfortunately, therefore, the results of this study support Andreasen and Manning's (1990) contention that Hispanic consumers may be vulnerable consumers who are not able to receive the same level of customer service afforded to other consumers in the United States market.

Implications for Consumer Affairs Managers

The data from this study should raise a warning flag for consumer affairs managers in the corporate world. There are abundant data, as presented earlier in this article, that clearly show that the Hispanic market is growing rapidly and many of these Hispanic consumers are uncomfortable speaking English. It appears that as many companies rush to market their products to this increasingly attractive Hispanic market, their customer service centers are not providing adequate support for these efforts. While many companies have service representatives who are fluent in Spanish, they have apparently neglected to analyze the entire process that these Spanish-language consumers must undertake to utilize fully these services. It is little wonder then that these customer service centers receive very few calls from Spanish-speaking consumers.

Without an adequate customer service telephone system in place for their Spanish-speaking consumers, consumer affairs managers run the considerable risk of failing to respond effectively to these consumers' questions and complaints. If dissatisfied Spanish-speaking consumers are unaware that they could call these toll-free customer service numbers and speak with a Spanish-speaking service representative, they are likely to engage in behaviors (i.e., negative word-of-mouth communication, switching to competitors' products) that have a detrimental impact on these companies. Therefore, the data in this study strongly indicate that this is a problem area that consumer affairs managers should remedy in order to better serve an important segment of their customers and thereby also benefit the

financial performance of their companies.

To encourage their Spanish-speaking consumers to communicate their questions or complaints to representatives in their service centers, consumer affairs managers should consider implementing a separate toll-free customer service number specifically targeted for use by Spanish-speaking consumers. On their product packages, companies could then list two separate toll-free customer service numbers, one for English-speaking consumers and another different number for Spanish-speaking consumers. Explicit directions should be written in Spanish on the packaging to encourage consumers to call this number if they have problems with the product and they would like to speak with a service representative who is fluent in Spanish. In addition, on this Spanish-language service line companies could also record directions in Spanish explaining to consumers how to wait on hold and how to use an automated service menu to route their calls. This approach, which should not be very expensive to implement, could greatly alleviate the problems that many Spanish-speaking consumers face when they attempt to use the customer service systems currently in place.

Public Policy Implications

From a strictly self-serving economic perspective, consumer affairs managers should be motivated to provide a better customer service telephone system for their Spanish-speaking consumers. However, if consumer affairs managers of companies are unwilling to respond positively to the needs of Spanish-speaking consumers, what role should the government play in encouraging or mandating appropriate corporate changes? Interestingly, this is an issue that is currently being addressed in a larger public policy context in the United States.

According to the 1990 U.S. Census, there are 323 different languages currently spoken in the United States (Mujica 1996). The 1990 Census also revealed that more than 31.8 million people in the United States speak a language other than English, which is a significant increase over the 23.1 million non-English speakers reported in the 1980 Census (Parlman and Shoeman 1994). In response to this increased use of languages other

than English, a movement has been launched in recent years to mandate the establishment of English as the official language of the United States (Lee 1997; Ojito 1997). Thus far, 23 states and a multitude of municipalities have passed laws declaring English to be the official language (Ojito 1997). In general, these laws require legislators to take the necessary steps to ensure that the role of English is preserved and enhanced (Parlman and Shoeman 1994).

Opinion polls show that these efforts to establish English as the official language have broad support. A Gallup poll found that 82 percent of respondents support English as the official language of the United States (O'Beirne 1996). At the same time, many employers in the United States have begun to require their foreign-born employees to converse only in English and not in their native languages (Ojito 1997; Parlman and Shoeman 1994). Unfortunately, some employers have taken this requirement to extremes by prohibiting their employees from speaking their native languages when they are on breaks, using the restrooms, or even walking outside the office buildings. As a result, advocacy groups are receiving record numbers of complaints from foreign-born workers regarding language discrimination in the workplace (Ojito 1997).

Against this backdrop, this question must then be raised "Should government require companies to provide better consumer service systems for their Spanish-speaking customers?" Proponents of stronger governmental mandates in this regard can readily point to the growing numbers of Spanish-speaking consumers in the United States. They may also draw upon the results of this study and argue that Spanish-speaking consumers are not currently served very well by the toll-free customer service systems that have been established thus far by consumer products companies.

However, from a pragmatic perspective, it appears that the time may not be ripe to consider the spawning of new governmental guidelines to rectify this defective customer service system for the benefit of Spanish-speaking consumers. Instead, the emergence and broad support for English-only laws indicates that the general public in the United States believes that residents in the United States who do not speak English should be

encouraged to learn to speak English. There is strong sentiment that immigrants should not be provided with additional support systems that allow them to continue to function in American society by communicating only in their native language. Thus, it is doubtful if regulations or legislation at the state or national level mandating better corporate customer service systems for Spanish-speaking consumers would receive much support.

CONCLUSION

The results from this study reveal that the toll-free customer service telephone systems currently used by consumer products companies do not address very well the needs of Spanish-speaking consumers. Even though many companies have Spanish-speaking representatives in their service centers, it is extremely doubtful that Spanish-speaking consumers could actually reach them to express their complaints and questions. A number of significant obstacles stand in the way of Spanish-speaking consumers. Most importantly, instructions on packages regarding the existence and use of toll-free customer service lines are written only in English, not in Spanish. Also, even if Spanish-speaking consumers do call these numbers, they are often directed in English to remain on hold for a service representative or to select a specific telephone button to direct their call to the proper destination.

Because of these obstacles, very few Spanish-speaking consumers currently call toll-free customer service centers. As a result, consumer affairs managers are not able to address the complaints and questions generated by this growing body of consumers in the United States. From a strictly economic perspective, consumer affairs managers should be motivated to correct this flaw in their current system so that they can more effectively meet the needs of this increasingly attractive market of Spanish-speaking consumers.

To rectify this weakness, consumer affairs managers should seriously consider adding separate toll-free customer service lines that are dedicated to Spanish-speaking consumers and staffed exclusively by Spanish-speaking service representatives. Consumer affairs managers should also provide information written in Spanish

on their products' packaging to encourage Spanish-speaking consumers to utilize this service. This new service option should remove most of the obstacles in the current system and thereby greatly increase the numbers of Spanish-speaking consumers who contact companies with their complaints and questions.

REFERENCES

- Albonetti, Joseph G. and Luis V. Dominguez (1989), "Major Influences on Consumer-Goods Marketers' Decision to Target U.S. Hispanics," *Journal of Advertising Research*, February/March, 9-21.
- Andreasen, Alan R. and Jean Manning (1990), "The Dissatisfaction and Complaining Behavior of Vulnerable Consumers," *Journal of Consumer Satisfaction/Dissatisfaction and Complaining Behavior*, 3, 12-20.
- Astor, David (1981), "The Hispanic Market: An In-Depth Profile," *Marketing Communications*, July, 15-19.
- Cornwell, T. Bettina, Alan David Bligh and Emin Babakus (1991), "Complaint Behavior of Mexican-American Consumers to a Third-Party Agency," *Journal of Consumer Affairs*, 25, 1, 1-18.
- Deshpande, Rohit, Wayne D. Hoyer and Naveen Donthu (1986), "The Intensity of Ethnic Affiliation: A Study of the Sociology of Hispanic Consumption," *Journal of Consumer Research*, 13, 2, September, 214-220.
- Diaz-Knauf, Katherine V., Howard G. Schutz and Roberta Y. Almeida (1992), "Hispanic Consumer Attitudes Toward Occupational Service Providers: Satisfaction, Quality, and Selection Criteria," *Journal of Consumer Satisfaction/Dissatisfaction and Complaining Behavior*, 5, 201-212.
- Fisher, Christy (1990), "Delivery Systems Foil Couponing," *Advertising Age*, 61, 43, October 15, 47.
- Fisher, Christy (1994), "Hispanics Indicate Enduring Preference for Native Language," *Advertising Age*, February 14, 26.
- Garrett, Dennis E. and Renee A. Meyers (1996), "Verbal Communication Between Complaining Consumers and Company Service Representatives," *The Journal of Consumer Affairs*, 30, 2, Winter, 444-475.
- Greene, Kenneth R. (1989), "The Social Context and Hispanics' Evaluations of Municipal Services," *Social Science Quarterly*, 70, 1, March, 205-210.
- Grimm, Cynthia J. and Dianne S. Ward (1997), "Competitive Edge or Over the Edge? Best Practices for Automated Response Systems," *Customer Relationship Management*, September, 5-7.
- Hamstra, Mark (1996), "Que Bueno! Chains Learn the Lucrative Language of Courting," *Nation's Restaurant News*, 30, 27, July 15, 67-72.
- Hitt, Greg (1998), "GOP Aims to Reverse Hispanic Shift Toward Democrats," *The Wall Street Journal*, May 14, A24.
- Kelly, J. Patrick, Jonathan C. Huefner and H. Keith Hunt (1991), "Frequency of Use and Level of Satisfaction with Rainchecks," *Journal of Consumer Satisfaction/Dissatisfaction and Complaining Behavior*, 4, 188-193.
- Lee, Tena Jamison (1997), "How Can There Be Free Speech If It's Only in English? The Debate over Language Rights," *Human Rights*, Spring, 24, 2, 10-17.
- Martin, Charles L. and Denise T. Smart (1994), "Consumer Experiences Calling Toll-Free Corporate Hotlines," *The Journal of Business Communications*, 31, 3, 195-212.
- Medina, Jose F., Joel Saegert and Alicia Gresham (1996), "Comparison of Mexican-American and Anglo-American Attitudes Toward Money," *Journal of Consumer Affairs*, 30, 1, Summer, 124-145.
- Millman, Joel (1998), "Mexican Food Brands Heat U.S. Market," *The Wall Street Journal*, April 15, A16.
- Moya, Steven (1985), "Foreign Relations: Implementing a Hispanic Marketing Program," *Sales & Marketing Management*, July 22, A-22-A-47.
- Mujica, Mauro (1996), "New Year's Resolution: U.S. English, Inc.," *National Review*, January 29, 48, 1, 46.
- Mulhern, Francis J. and Jerome D. Williams (1994), "A Comparative Analysis of Shopping Behavior in Hispanic and Non-Hispanic Market Areas," *Journal of Retailing*, 70, Fall, 231-251.
- Murphy, Ian P. (1997), "Marketing Briefs," *Marketing News*, 31, 5, March 3, 2.
- Nicholls, J. A. F., Sydney Roslow and Sandipa Dubish (1997), "Hispanic and Non-Hispanic Mall Shoppers: Segmentation by Situational Variables and Purchase Behaviors," *Journal of Segmentation in Marketing*, 1, 1, 57-73.
- O'Beirne, Kate (1996), "Bread & Circuses," *National Review*, July 1, 48, 12, 21.
- O'Guinn, Thomas C. and Timothy P. Meyer (1983), "Segmenting the Hispanic Market: The Use of Spanish-Language Radio," *Journal of Advertising Research*, 23, 6, December, 9-16.
- Ojito, Mirta (1997), "Bias Suits Increase Over English-Only Rules," *The New York Times*, April 23, 146, B1-B2.
- Parlman, Gregory C. and Rosalie J. Shoeman (1994), "National Origin Discrimination or Employer Prerogative? An Analysis of Language Rights in the Workplace," *Employee Relations Law Journal*, 19, 4, Spring, 551-565.
- Petrozzello, Donna (1995), "Radio Targets Hispanic Niches," *Broadcasting & Cable*, November 13, 76.
- Saegert, Joel, Robert J. Hoover and Marye Tharp Hilger (1985), "Characteristics of Mexican American Consumers," *Journal of Consumer Research*, 12, 1, June, 104-109.

- Segal, Madhav N. and Lionel Sosa (1983), "Marketing to the Hispanic Community," *California Management Review*, XXVI, 1, Fall, 120-134.
- Shim, Soyeon and Kenneth C. Gehrt (1996), "Hispanic and Native American Adolescents: An Exploratory Study of Their Approach to Shopping," *Journal of Retailing*, 72, 3, 307-324.
- SOCAP (Society of Consumer Affairs Professionals in Business) (1992), *SOCAP 800 Number Study: A 1992 Profile of 800 Numbers for Customer Service*, Alexandria, VA: Society of Consumer Affairs Professionals in Business.
- Valdes, M. Isabel and Marta H. Seoane (1995), *Hispanic Market Handbook*, New York: Gale Research Inc.
- Valencia, Humberto (1983), "Point of View: Avoiding Hispanic Marketing Blunders," *Journal of Advertising Research*, 23, 6, 19-22.
- Valencia, Humberto (1989), "Hispanic Values and Subcultural Research," *Journal of the Academy of Marketing Science*, 17, 1, 23-28.
- Vargas, Alexia (1997), "Harvard MBAs' Chain Riles Chicago's Hispanic Grocers," *The Wall Street Journal*, November 6, B1.
- Villareal-Camacho, Angelina (1983), "Consumer Complaining Behavior: A Cross-Cultural Comparison," *1983 American Marketing Association Educators Conference Proceedings*, Patrick E. Murphy et al., (eds.), 68-73.
- Wallendorf, Melanie and Michael D. Reilly (1983), "Ethnic Migration, Assimilation, and Consumption," *Journal of Consumer Research*, 10, 3, December, 292-302.
- Wynter, Leon E. (1997a), "Blacks and Hispanics Gain Spending Clout," *The Wall Street Journal*, September 3, B1.
- Wynter, Leon E. (1997b), "Hispanic Buying Habits Become More Diverse," *The Wall Street Journal*, January 8, B1.
- Zbar, Jeffrey D. (1995), "Dynamic Hispanic Audience Expands," *Advertising Age*, February 6, 35.

Send correspondence regarding this article to:

Dennis E. Garrett
Marketing Dept., 432 Straz Hall
Marquette University
P.O. Box 1881
Milwaukee, WI 53201-1881 USA

COMPARATIVE ANALYSIS OF DISSATISFIED VERSUS SATISFIED HIGHER INCOME HISPANIC CATALOG CONSUMERS: SHOPPING ORIENTATION AND STORE ATTRIBUTES

Marianne C. Bickle, Colorado State University
Molly Eckman, Colorado State University
Antigone Kotsiopoulos, Colorado State University

ABSTRACT

A comparison of dissatisfied and satisfied higher income Hispanic catalog consumers' shopping orientations and the influence store attributes had on store choice was made. The sample consisted of 178 higher income Hispanics living in Los Angeles, California, in San Antonio, Texas, and in New York City, New York. Multivariate and univariate analyses of variance were used to examine differences between the groups. Results of the analyses revealed satisfied higher income Hispanic catalog consumers to be significantly more prone to home shopping than dissatisfied consumers. Regardless of their professed level of satisfaction with catalog shopping, neither group shopped through catalogs on a frequent basis. Dissatisfied and satisfied higher income Hispanics believed merchandise offering, bilingual information, customer service, and value for price influenced store choice. Many catalog retailers have offered these store attributes at the same level or superior to that of in-store retailers. Despite such efforts, catalog retailers have had limited success in building patronage loyalty among higher income Hispanic consumers.

INTRODUCTION

Seven percent of the approximately 2.9 million Hispanic households in the United States have been classified as higher income (Braun, 1991). During 1993, higher income Hispanics' average annual income was \$33,000, approximately \$10,000 more than Hispanics in general. Due to their discretionary income and access to credit cards (Wynter, 1995), higher income Hispanics have been identified as a potentially profitable target market. Unfortunately, this group has consistently been neglected in research. To date, only one published study has examined higher income Hispanics (i.e., Eckman, Bickle and Kotsiopoulos, 1997). The purpose of this study was to reveal

preliminary shopping data on higher income Hispanics.

Catalog retail sales have continued to grow at an annual rate of 6.1% (Direct Marketing Association, 1996). Despite such increasing popular acceptance of catalogers, higher income Hispanic consumers in Colorado shopped significantly less through catalog or mail order than did non-Hispanics (Eckman et al., 1997). The deficiency of empirical research on higher income Hispanics' consumer behavior has left marketers and retailers without an accurate understanding of where this group shops and what they purchase. Understanding higher income Hispanic consumers' dissatisfaction/satisfaction may help marketers understand their catalog-shopping behavior. This understanding could lead to enhanced marketing strategies to entice higher income consumers to shop by mail.

Catalog Retailing

During 1996, catalog retail sales comprised 60% of all direct marketing sales (Direct Marketing Association, 1996). The more than 12 billion catalogs mailed each year target over 100 million U.S. adults annually (Sroge, 1995). Consumers spent over \$40 billion on merchandise sold through catalogs during 1996. Annual purchases are expected to rise to \$51 billion by the year 2000 (Direct Marketing Association, 1996). Why have higher income Hispanics not readily adopted the convenience of catalog shopping? Would an analysis of their shopping orientations reveal a specific enjoyment of patronizing in-store retailers? Are catalog store attributes contrary to higher income Hispanics' needs?

Consumer acceptance of catalog retailing has been inextricably tied to a diverse range of factors within the changing marketplace. These factors include a changing demographic profile, increased size of the elderly population, decreased discretionary time and increased number of women

in the work force. Store attributes such as brand name merchandise, 1-800 numbers, 24-hour purchasing capability, lower risks associated with catalog shopping than previously experienced, and the adoption of consumer-friendly services have promoted catalog purchases (Lucas, Bush, and Gresham, 1994; Simone, 1992; Sroge, 1995).

Businesses have relied on the analysis of demographics to provide valuable insight into consumer patterns (Eastlick, 1994; Shim and Bickle, 1993). For example, research has revealed the typical catalog shopper to be female (Jasper and Lan, 1992), married (Loudon and Della Bitta, 1988) and better educated. They also earn a higher than average income and are non-Hispanic (Eckman et al., 1997). When catalog texts were printed in English and Spanish, direct marketers often realized a greater than average response rate from Hispanic consumers (Reynardus, 1992). The significant increase in catalog retailers and higher income Hispanic consumers warrant obtaining a better understanding of this market.

Due to the competition among the 12,000 catalogers in the U.S. consumers have been selective concerning whom they patronized, how often they purchased, and how many of the products purchased were kept. Clearly businesses must satisfy consumers' needs or risk losing valuable and potentially loyal consumers (Sroge, 1995). At a time when catalog shopping is at its peak of popularity, why have higher income Hispanics resisted catalog shopping? Why has this group been so elusive to marketers and retailers?

Shopping Orientation

In conjunction with the demographic analysis, understanding consumers' shopping orientation has assisted retailers in predicting the target market's shopping behavior (Bickle and Shim, 1993; Gehrt, Yale and Lawson, 1996). The image obtained may be used to more accurately assess consumers' shopping priorities and patronage motives (Jasper and Lan, 1992; Miller, 1993). Based on shopping orientations, catalog consumers have been profiled as cosmopolitan, style and value conscious, and convenience-oriented (Kwon, Soea and Arzeni, 1991). They frequently purchased products using credit cards, were less conservative (Erevelles and Leavitt, 1992), were relatively unconcerned with

risks typically associated with catalog purchases and read catalogs front to back (Direct Marketing Association, 1996). The sole data published on higher income Hispanics is that they are not frequent catalog consumers (Eckman et al., 1997). Research may assist direct marketers understanding of the attributes that motivate this group catalog patronage. Without empirical research as a foundation for actions, marketing decisions will be based primarily on stereotypes.

Store Attributes

Store attributes have been found to be influential on store choice. Eckman et al. (1997) found that language spoken, pricing policy, store hours, comfort, and merchandise selection were important to higher income Hispanic consumers. A logical next step is to examine level of satisfaction with these attributes. If catalog retailers are to attract an increasing portion of the higher income Hispanic market, an analysis of their preference for or satisfaction with store attributes is needed.

Consumer Satisfaction

In an attempt to increase consumers' satisfaction, catalog retailers have improved and expanded the services offered (Sroge, 1995). Rewards for services offered have included increased market share, customer store loyalty, repeat purchase behavior and positive word-of-mouth promotion (Erevelles and Leavitt, 1992). Satisfying higher income Hispanic consumers has been more difficult than anticipated. Problems in attracting and satisfying this market have been due in part to forced segmentation by marketers, inaccurate use of the Spanish language and ignoring ethnic values (e.g., family, culture) (Fisher, 1991; Zuckerman, 1990). Understanding the perceptions and needs of dissatisfied and satisfied higher income Hispanic catalog consumers is one essential component toward learning more about this lucrative market.

Conceptual Framework

An adapted version of the Monroe and Gultinan (1975) model was the conceptual

framework for this study, focusing on components of the model that 1) affected shopping orientations, 2) influenced perceptions of store attributes and 3) influenced attitudes toward stores, which have influenced store choice. The model was adapted for this study by including ethnicity (i.e., higher income Hispanics) in consumer characteristics and satisfaction with catalog retailers. Despite the buying potential of the higher income Hispanic market, research focusing on level of satisfaction with catalog retailers has not been identified. The analysis may advance the theoretical base of the Monroe and Gultinan framework (1975). Practical application of the results may provide retailers with insight regarding the level of satisfaction with catalogs as they relate to shopping behaviors.

METHOD

Sample

The sample consisted of higher income Hispanic consumers living in Los Angeles, California, in San Antonio, Texas and in New York City, New York. These cities were selected for the study due to the high number of Hispanics living in each area. Sampling criteria included consumers 18 years and older with an annual household income equal to or higher than the national mean of \$32,000 per year (U.S. Bureau of the Census, 1990). A national mailing list firm utilized a systematic random sampling method to identify 3,000 higher income Hispanic subjects meeting the criteria. The sample consisted of 299 useable questionnaires (10% response rate). This sample was part of a larger database.

Instrument

A Spanish and English version of the questionnaire was developed using the back translation method. Sections of questions included: demographics, shopping orientation, influence of store attributes on store choice and level of satisfaction with catalog shopping.

Demographics. A series of demographic questions were posed. Data analyzed included respondents' age, education, income, marital

status, occupation, gender and country of origin.

Shopping orientation. Previous research on consumers' shopping orientation was the foundation for this section of the study (e.g., Bickle and Shim, 1993; 1994). Thirty-nine shopping orientation statements were included in the survey and measured on a 5-point Likert-type scale (1 = strongly disagree; 3 = neutral; 5 = strongly agree). Principal component factor analysis with varimax rotation was conducted. Statements loading greater than .40 were included in the preliminary analysis (see table 1). Thirteen factors with an eigenvalue greater than 1.0 were generated (see table 2). Cronbach alpha coefficients were conducted on each factor. Factors with an alpha coefficient of .70 or higher were used in the analysis (these criterion are used throughout the study). The three shopping orientation factors included in the analysis were labeled 1) insecure shopper, 2) credit card user and 3) home shopper (see table 3).

Store attributes. The influence of 45 store attributes on patronage behavior was measured on a 5-point Likert-type scale (1 = not at all influential; 3 = neutral; 5 = very influential). The majority of attributes were based on previous research (e.g., Shim and Mahoney, 1991). The remaining statements were developed for this study. Principal component factor analysis was conducted on the statements (see table 4). Eleven factors with an eigenvalue greater than 1.0 were generated (see table 5). Seven store attribute factors had a Cronbach alpha coefficient of .70 or greater. These factors were labeled 1) bilingual, 2) services, 3) customer service, 4) extended store hours, 5) value for price, 6) merchandise offering and 7) ambiance (see Table 6).

Satisfaction. Higher income Hispanic consumers' level of satisfaction with catalog retailers was the focus of this study. Building upon previous research (e.g., Bickle, Kotsiopulos, Dallas and Eckman, 1995), a 5-point Likert-type scale was used to measure consumers' satisfaction with catalog retailers (1 = very dissatisfied; 3 = neutral; 5 = very satisfied). Consumers were segmented into two groups - - dissatisfied versus satisfied catalog consumers. Respondents who

Table 1
Factor Analysis: Shopping Orientation Statements

Factor	1	2	3	4	5	6	7	8	9	10	11	12	13
Question													
1	-.004	.212	-.128	-.061	.168	.044	.165	.354	-.010	.649	.130	.049	-.046
2	.035	.086	.133	.055	.054	-.020	.026	.863	.011	-.040	.050	-.048	.017
3	-.034	.136	-.026	.046	.049	.002	.011	.852	-.000	.068	.147	.081	.086
4	.052	-.029	-.067	-.109	.121	.210	-.006	.199	.081	-.288	.645	.026	-.073
5	.115	.046	.874	-.006	.010	.095	-.017	.070	.062	-.080	.087	.003	-.137
6	.776	.084	.098	.179	-.068	-.098	.012	.033	.074	-.077	.104	-.033	.051
7	.039	.117	.189	-.013	-.034	-.032	.099	.101	.017	.249	.728	.004	.009
8	.070	.014	.013	-.118	.295	-.035	.680	.090	-.016	.209	-.088	.205	-.074
9	.035	.021	.043	-.106	.148	-.069	.117	.128	.083	-.044	-.072	.034	.735
10	.139	.151	.172	.405	.144	-.011	.086	.167	.267	.047	-.146	-.109	-.447
11	.042	-.020	.015	.240	-.014	.216	.241	-.044	.563	.341	-.140	.078	.055
12	.004	-.005	.069	-.052	.046	-.027	.046	.051	.803	-.082	.072	.123	-.032
13	.126	.138	.090	-.010	.028	.703	-.039	.024	.193	-.037	-.116	.053	-.152
14	.059	.242	.015	.089	.167	.025	.752	.065	.009	-.003	-.011	-.026	.019
15	-.164	-.057	-.009	-.123	.105	-.114	.036	-.121	.052	.724	-.047	-.080	-.056
16	.098	.702	.033	-.068	-.181	.175	-.034	.152	.076	.114	-.042	-.038	.124
17	.833	.090	.055	.105	.058	-.024	.006	-.107	.056	-.006	-.089	.008	-.016
18	.205	-.001	-.019	.697	.198	-.001	-.095	-.040	.081	-.022	-.138	.061	-.104
19	.088	.070	-.062	.256	.251	-.021	.041	-.055	.389	-.238	.293	-.109	.332
20	.044	.748	.011	.083	.091	-.015	.151	.024	.099	-.177	-.070	-.007	-.037
21	-.035	.534	-.066	.018	.511	-.102	.036	.039	.020	.174	.152	.060	-.068
22	.153	-.007	.021	.235	.621	.216	.018	.213	.030	.275	-.072	-.007	.209
23	.029	-.035	.079	.121	.738	-.038	.202	-.003	.104	.018	.068	.179	.065
24	.117	.438	.528	.160	.174	-.129	.046	-.061	-.083	.135	.146	-.005	.034
25	.704	.010	.082	.026	.175	.172	.104	.082	-.022	.044	.142	.089	.025
26	.191	-.042	-.083	.540	-.015	.054	.029	-.000	-.128	-.106	.197	-.051	-.132
27	.061	.245	.061	-.116	.081	.321	.064	-.148	-.175	.002	.104	.387	-.316
28	.046	.028	.067	.038	-.075	.050	.752	-.063	.079	.017	.154	.034	.126
29	-.019	.320	.154	.275	-.180	.578	.115	-.024	-.209	.209	.070	.049	.090
30	.015	-.035	.015	.012	.456	.547	.199	.043	.222	.012	-.040	-.192	-.006
31	-.089	-.131	-.029	.047	.000	.648	-.038	-.035	-.023	-.171	.326	.085	.037
32	.094	.060	.861	-.027	-.010	.083	.087	.053	-.017	-.042	-.038	.077	.114
33	.800	-.000	.042	.064	-.021	.026	.048	.014	.010	-.104	-.050	.088	-.038
34	.041	.651	.223	.147	-.029	.067	.147	.142	-.023	.050	.115	.015	-.097
35	.034	.117	.007	.567	.163	.088	.178	.012	.211	.157	-.212	.159	.199
36	.004	.182	.123	.506	-.014	.024	.004	.170	-.096	-.179	.018	.179	-.006
37	.125	-.022	.010	.091	.310	-.023	.111	.147	.177	-.054	-.102	.692	-.113
38	.055	-.050	.081	.235	-.102	.075	.071	-.057	.248	-.006	.100	.701	.270
39	.094	.209	-.116	-.088	.147	.134	-.121	-.035	.458	.102	.047	.164	.110

were very dissatisfied or dissatisfied with catalog shopping comprised the dissatisfied group (n = 59). Respondents who were satisfied or very satisfied made up the satisfied group (n = 119). Respondents who expressed a neutral attitude toward catalog retailers were not included in the analysis (n=121).

HYPOTHESES

H1: Significant differences exist between dissatisfied and satisfied higher income Hispanic catalog consumers' shopping orientations.

H2: Significant differences exist between the perceived influence store attributes had on

dissatisfied and satisfied higher income Hispanic catalog consumers' store choice.

Table 2
Preliminary Shopping Orientation Factors

Factors	Eigenvalue	Percent of Variance	Cumulative Percent
1	5.16	13.2	13.2
2	2.77	7.1	20.4
3	2.54	6.5	26.9
4	1.99	5.1	32.0
5	1.87	4.8	36.8
6	1.73	4.5	41.5
7	1.58	4.1	45.3
8	1.44	3.7	49.0
9	1.36	3.5	52.5
10	1.22	3.1	55.7
11	1.17	3.0	58.7
12	1.07	2.8	61.4
13	1.05	2.7	64.1

Table 3
Principal Component Factor Analysis:
Shopping Orientation Statements

Factor label and statements	Chronbach Alpha Coefficient
Insecure Shopper ◆ I often seek the advice of friends regarding which brands to buy. ◆ I often seek the advice of family members regarding which brands to buy.	.81
Credit Card User ◆ I buy many things with a credit card. ◆ I use credit cards because they offer me clout. ◆ Paying with a credit card is more convenient than paying with cash.	.73
Home Shopper ◆ I usually have more success shopping by mail, phone or computer shopping in the store. ◆ I find catalog shopping convenient. ◆ I would use home TV shopping. ◆ Shopping at home is more efficient than shopping at a store.	.70

RESULTS

Respondents' Profile

Respondents tended to be between the ages of

19 and 50 (95%), female (54%) and married (64%). Forty eight percent of respondents reported having attended college. The majority held full-time (74%) professional positions (54%) and earned between \$32,000 and \$50,000 (88%). Respondents' national origin included Cuba (48%), Mexico (22%) and Puerto Rico (30%).

Shopping Orientation

Multivariate analyses of variance (MANOVA) revealed an overall difference between dissatisfied and satisfied higher income Hispanic consumers' shopping orientation (Multivariate $F(1,85) = 12.86, p < .001$). Partial support for Hypothesis 1 was supported. Univariate analysis of variance indicated satisfied higher income Hispanic catalog consumers were home shoppers to a greater extent than dissatisfied Hispanic catalog consumers ($F(1,176) = 45.48, p < .001$). Despite the significant difference in level of satisfaction, dissatisfied ($M=1.83$) and satisfied ($M=2.78$) higher income Hispanics did not perceive themselves as home shoppers. MANOVA failed to reveal significant differences between groups' shopping orientation as a credit card user (see table 7). Dissatisfied ($M = 2.90$) and satisfied ($M = 2.94$) higher income Hispanics expressed neutral attitudes toward the use of credit cards. Both groups also were neutral about being insecure shoppers (dissatisfied shoppers, $M = 2.73$; satisfied consumers, $M = 2.48$).

Store Attributes

Multivariate analysis of variance failed to reveal an overall difference between the influence store attributes had on higher income Hispanics' decision to shop at a particular store. Support for Hypothesis 2 was not provided (see table 8). Dissatisfied ($M = 4.15$) and satisfied ($M = 4.18$) higher income Hispanic catalog shoppers were positively influenced by a store's level of customer service. Dissatisfied ($M = 4.39$) and satisfied ($M = 4.15$) higher income Hispanic consumers perceived a store's ambiance as being influential on the outcome of which store to chose. Value for price was identified as an important store attribute by dissatisfied ($M = 4.16$) and satisfied ($M = 4.30$) consumers. Both groups felt neutral or

Table 4
Factor Analysis: Influence of Store Attributes on Store Choice

Factor	1	2	3	4	5	6	7	8	9	10	11
Question											
1	.167	-.012	.141	.238	.106	-.096	-.048	.202	.022	.087	.761
2	-.067	.163	-.054	.177	-.012	-.027	-.077	.837	-.012	-.001	.110
3	-.099	.084	.007	.239	.004	.031	.071	.805	.656	-.084	.046
4	-.078	.127	.075	.688	-.004	-.031	-.098	.291	.159	-.023	-.096
5	.095	-.043	.002	.652	.088	.145	.190	.040	.159	.000	.135
6	.119	.068	.893	.485	-.020	.014	.135	.007	.054	.088	.077
7	.242	-.030	.564	.111	.362	.018	.057	.005	-.161	.000	-.096
8	.097	.065	.924	.007	.016	.055	.045	.024	.031	.000	.055
9	.115	.071	.929	.020	-.042	.070	.101	-.002	.032	.021	.047
10	-.143	.289	.229	.302	-.077	.114	.103	-.091	-.083	.496	.262
11	.092	.197	.031	.700	-.046	.039	.064	.420	.033	.124	.107
12	.174	.375	-.053	.433	.058	-.114	.152	.255	-.107	.295	-.341
13	.073	.154	.048	.790	.070	.169	-.081	.032	-.089	-.055	.065
14	.632	.161	.126	.070	-.101	.332	.022	.013	-.163	-.249	-.018
15	.628	.170	.107	.216	.157	.077	.013	-.054	-.067	-.404	-.029
16	.159	.443	.006	.329	.183	.125	.135	.096	.113	-.542	.013
17	.228	.130	.258	.130	.032	-.072	.666	-.078	.117	-.177	-.042
18	.072	.781	.022	.132	.053	-.063	.090	.049	.109	-.128	.034
19	.143	.511	.082	.238	.123	.273	.276	.128	-.033	-.044	-.060
20	.182	.542	.143	.011	-.023	.231	.230	.157	.243	.062	.066
21	.135	.406	.106	.083	.186	.358	.454	.007	-.068	-.017	-.128
22	.123	.270	.187	-.029	.122	.326	.440	.280	-.170	.044	.212
23	.199	.225	.083	.007	.264	.142	.470	.394	-.155	.164	-.042
24	.114	.809	-.012	.034	.079	.126	.126	-.044	-.004	-.028	.088
25	.165	.826	.080	.101	.100	.005	-.009	.101	.118	.065	.015
26	.221	.771	.035	.117	.066	.066	-.114	.151	.213	.074	-.078
27	.385	.405	.181	.207	.056	.122	-.121	.279	.293	.135	-.257
28	.546	.088	.189	-.023	.296	.157	.163	-.071	.081	.360	-.099
29	.648	.069	.149	.106	.390	-.025	.137	-.027	.083	-.015	-.034
30	.734	.171	-.000	.096	.041	.045	.155	.019	.039	.048	.184
31	.637	.157	.136	.022	-.010	.118	.248	-.059	.217	.127	.144
32	.615	.075	.196	.100	.166	-.016	.036	-.121	.185	-.239	-.134
33	.720	.233	.054	-.064	.054	.349	.043	.032	.019	.002	.028
34	.654	.111	.027	-.070	.131	.097	.216	.044	.254	.143	.125
35	.355	.188	.071	.044	.679	-.111	-.005	-.003	-.140	.116	.087
36	.160	.183	-.043	.044	.742	.182	.086	.011	.247	-.165	.048
37	.016	.036	.018	.005	.701	.295	.172	.047	.304	-.093	.011
38	.173	.256	.012	.092	.243	.224	-.008	.083	.720	-.002	.085
39	.245	.297	.024	.117	.183	.094	.175	-.085	.619	-.130	-.059
40	.418	-.016	.118	-.040	.100	.166	.650	-.098	.230	.123	-.017
41	.340	.066	.112	.170	.077	.605	.218	.010	.218	-.102	-.013
42	.219	.175	.071	.155	.162	.075	.100	.022	.123	.026	-.074
43	.171	.366	-.086	.137	.108	.052	.106	.045	.394	.146	-.001

Table 5
Preliminary Store Attribute Factors

Factors	Eigenvalue	Percent of Variance	Cumulative Percent
1	11.90	26.5	26.5
2	3.79	8.4	34.9
3	2.95	6.6	41.5
4	2.04	4.6	46.0
5	1.74	3.9	49.9
6	1.69	3.8	53.7
7	1.53	3.4	57.1
8	1.39	3.1	60.2
10	1.13	2.5	65.4
9	1.20	2.7	62.8
11	1.01	2.3	67.6

Table 6
Principal component factor analysis: Store characteristics that influence where you shopped

Factor label and sample statements	Chronbach Alpha Coefficient
Bilingual	.93
◆ Spanish speaking sales associates	
◆ Advertising and signs in Spanish	
◆ Bilingual information (3)	
Services	.89
◆ Toll free telephone shopping	
◆ Packaging and gift wrapping	
◆ Information mailer (9)	
Customer Services	.87
◆ Product knowledge of sales personnel	
◆ Easy complaint procedure	
◆ Convenient return policy	
Extended Store Hours	.78
◆ Late evening store hours	
◆ Sunday store hours (2)	
Value for Price	.76
◆ Frequency of sales	
◆ Selection of price ranges	
◆ Discounts or special sales (4)	
Merchandise Offering	.75
◆ Merchandise quality	
◆ A variety of merchandise selection (2)	
Ambiance	.70
◆ Familiarity with store and sales associates	
◆ Attractive décor (4)	

() Represents the number of statements in the factor

disagreed that the store's services, extended store

hours, and merchandise offerings had little influence when selecting a store.

DISCUSSION, LIMITATIONS, AND RECOMMENDATIONS

The large size of higher income Hispanic consumers and their significant purchasing power (Winsberg, 1994) and large number of catalog retailers provided sound justification for examining higher income Hispanics' shopping orientations and influence store attributes have made on store choice. Satisfied higher income Hispanic catalog consumers were shown to be significantly more prone to shopping at home than dissatisfied higher income Hispanic catalog consumers. However, the mean scores revealed that despite the group's satisfaction with catalog retailers, higher income Hispanics did not consider themselves to be home shoppers. They also have not regularly used credit cards. Due in part to the nature of payment (i.e., pay by mail or phone when using a credit card), catalog shoppers have predominately used a credit card as the preferred mode of payment (Erevelles and Leavitt, 1992). Higher income Hispanics' preference to pay by cash or check may have significantly influenced their home shopping orientation.

Dissatisfied and satisfied higher income Hispanics were not perceived as insecure shoppers. This finding suggests that risks typically associated with catalog shopping (e.g., unable to try on garments prior to the purchase) may not have negatively influenced their decision to patronize catalog retailers. As infrequent catalog shoppers, higher income Hispanics may have misunderstood the quality and variety of attributes offered (e.g., variety of product offerings, ease of ordering) (Eckman et al., 1997). They may also require additional information on attributes deemed as important for patronage. An in-depth examination of important attributes may assist retailers in reaching this yet untapped market.

Customer service offered by catalogers has been identified as being superior to that of many traditional retailers. Unlike many in-store retailers, catalog retailers have expanded the number and improved the quality of services (Collinger, 1995; Sroge, 1995). Higher income Hispanics may not have recognized the services

Table 7
Multivariate and Univariate Analysis of Variance: Shopping Orientations of Higher Income Dissatisfied and Satisfied Hispanic Consumers

Shopping Orientation Factors	Mean Score (S.D.)		Univariate F	Multivariate F
	Dissatisfied Consumers	Satisfied Consumers		
Insecure Shopper	2.73 (1.08)	2.48 (1.10)	2.24	12.86***
Credit Card User	2.90 (1.15)	2.94 (1.09)	.06	
Home Shopper	1.83 (.71)	2.78 (.96)	45.48***	

***p < .001

Table 8
Multivariate and Univariate Analysis of Variance: Influence of Store Attributes on Higher Income Hispanic Catalog Shoppers Level of Satisfaction

Store attribute factors	Means (S.D.)		Univariate F	Multivariate F
	Dissatisfied Catalog Shoppers	Satisfied Catalog Shoppers		
Bilingual	4.28 (.92)	4.15 (1.08)	.41	1.32
Service	3.15 (1.05)	3.34 (.95)	.27	
Customer Service	4.15 (.70)	4.18 (.80)	.79	
Extended Store Hours	2.46 (1.26)	2.11 (1.27)	.41	
Value for Price	4.16 (.76)	4.30 (.76)	.28	
Merchandise Offerings	4.43 (.58)	4.46 (.82)	.83	
Ambiance	3.20 (.94)	3.20 (.98)	.93	

offered due to infrequent use of catalogs. An emphasis on the significant changes made by catalog retailers may alter higher income

Hispanics' appreciation and patronage of catalogs.

Sound actions by catalog retailers' demand

additional research on this important yet little known consumer group. The population size, disposable income and preferences for store attributes have made higher income Hispanic consumers an important part of catalog retailers' future growth. As the next viable market to serve, additional information on this group is warranted. Focus groups may provide insight on higher income Hispanics' infrequent patronage of catalog retailers and preference for in-store retailers. Important store attributes ignored in research may also be revealed.

REFERENCES

- Bickle, M. C., A. Kotsiopoulos, M. J. Dallas and M. Eckman (1995), "Fit of Women's Jeans: An Exploratory Study Using Disconfirmation Paradigm," *Journal of Consumer Satisfaction/Dissatisfaction and Complaining Behavior*, 8, 208-213.
- Bickle, M. C. and S. Shim (1994), "Customers' Overall Satisfaction with Services by Retailers Type: A Function of Individual Characteristics and Expected Ideal Level of Service," *Journal of Consumer Satisfaction/Dissatisfaction and Complaining Behavior*, 7, 198-206.
- Bickle, M. C. and S. Shim (1993), "Usage Rate Segmentation of the Electronic Shopper: Heavy Versus Non-Heavy Spenders," *International Review of Retail, Distribution and Consumer Research*, 3, (1), 1-18.
- Braun, H. D. (1991), "Marketing to Minority Consumers," *Discount Merchandiser*, 31, (February), 45, 75.
- Collinger, T. (1995), "What Direct Agencies Will Need to Succeed," in *Beyond 2000: The Future of Direct Marketing*, J. I. Reitman, (ed.), NTC Business Books: Lincolnwood, Illinois.
- Direct Marketing Association (1996), *Statistical Fact Book*, New York: Direct Marketing Association.
- Eastlick, M. A. (1994), "How Many Catalogs are too Many?" *Stores*, 76, (1), (January), R5-RR7.
- Eckman, M., A. Kotsiopoulos and M. Bickle (1997), "Store Patronage of Hispanic versus Non-Hispanic Consumers: Comparative Analyses of Demographics, Psychographics, Store Attributes, and Information Search," *Hispanic Journal of Behavioral Sciences*, 19, (1), 69-83.
- Erevelles, S. and C. Leavitt (1992), "A Comparison of Current Models of Consumer Satisfaction/Dissatisfaction," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 5, 104-114.
- Fisher, C. (1991), "Study: Separate Ad Buys Needed to Reach Hispanics," *Advertising Age*, 62, (18), (April 29), 54.
- Gehrt, K. C., L. J. Yale and D. A. Lawson (1996), "The Convenience of Catalog Shopping: Is There More to it Than Time?" *Journal of Direct Marketing*, 10, (Autumn), 19-28.
- Jasper, C. R. and P. R. Lan (1992), "Apparel Catalog Patronage: Demographic, Lifestyle and Motivational Factors," *Psychology & Marketing*, 9, (July/August), 275-296.
- Kwon, Y-H, S. L. Soae and M. Arzeni (1991), "Catalog vs. Non-Catalog Shoppers of Apparel: Perceived Risks, Shopping Orientations, Demographics, and Motivation," *Clothing and Textile Research Journal*, 10 (Fall), 13-19.
- Loudon, D. and A. J. Della Bitta (1988), *Consumer Behavior: Concepts and Applications*, New York: McGraw-Hill.
- Lucas, G. H., Jr., R. P. Bush and L. G. Gresham (1994), *Retailing*, Houghton Mifflin: Boston, MA.
- Miller, C. (1993), "Study: Shopping Patterns Vary Widely Among Minorities," *Marketing News*, 27, (2), (January 18), 11.
- Monroe, K. B. and J. P. Guiltinan (1975), "A Path-Analytic Exploration of Retail Patronage Influences," *Journal of Consumer Research*, 2, 19-28.
- Reynardus, J. E. (1992), "Tips on Targeting Hispanics Through Direct Marketing," *Potential in Marketing*, 25, (5), (June), 18.
- Shim, S. and M. Bickle (1993), "Women 55 Years and Older as Catalog Shoppers: Satisfaction with Apparel Fit and Catalog Attributes," *Clothing and Textiles Research Journal*, 11, (3), 53-64.
- Shim, S. and M. Mahoney (1991), "Electronic Shoppers and Non-Shoppers Among Videotex Users: Shopping Orientations, Videotex Usages and Demographics," *Journal of Direct Marketing*, 5, (3), 29-38.
- Simone, J. (1992), "The Mailing of America," *VM&SD*, (November), 24-32.
- Sroge, M. (1995), "The Future of Catalogs: New Technology, New Markets, New Horizons," *Beyond 2000: The Future of Direct Marketing*, J. I. Reitman, (ed.), NTC Business Books: Lincolnwood, Illinois.
- U.S. Bureau of the Census (1990), "Summary of Population and Housing Characteristics," *U.S. Bureau of the Census*, Washington, DC: U.S. Government Printing Office.
- Winsberg, M. (1994), "Specific Hispanics," *American Demographics*, 16, (2), (February), 4-53.
- Wynter, L. E. (1995), "Reaching Hispanics Across a Racial and Cultural Divide," *Wall Street Journal*, (February 15), B1.
- Zuckerman, D. (1990), "The Hispanic Market," *Restaurant Business Market*, 89, (1), (March 1), 89.

Send correspondence regarding this article to:

Marianne C. Bickle
 Dept. of Design, Merchandising and Consumer Sciences
 161 Aylesworth
 Colorado State University
 Fort Collins, CO 80523-1575 USA

EFFECT OF EXPECTATIONS ON USER SATISFACTION WITH A REMODELED UNIVERSITY DINING FACILITY

Carol C. Caughey, Oregon State University
Sally K Francis, Oregon State University
Anka E. Kolodziej, Oregon State University

ABSTRACT

The effect of user expectations on satisfaction with the interior of a recently remodeled student owned dining facility was explored. The dependent variable was overall satisfaction. The independent variable, expectations, was operationalized as whether or not users had seen the facility prior to the time it was closed for remodeling. An analysis of covariance was conducted using major as a covariable. As suggested by disconfirmation theory, expectations were found to have had a significant effect on satisfaction. The users who had seen the facility prior to the remodel reported significantly lower overall satisfaction with the remodeled facility than those who had not.

INTRODUCTION

The theoretical framework for the present study was based on disconfirmation of expectations. Research on satisfaction using disconfirmation of expectations suggests that satisfaction is the result of a comparison of that which was expected with that which was received (Woodruff, Cadotte, & Jenkins, 1983). An underlying premise of disconfirmation of expectations is that expectations are related to satisfaction. Postpurchase evaluation of a product can be explained, at least in part, by a comparison of the prepurchase expectations and the evaluation of the postpurchase performance (Erevelles & Leavitt, 1992). In a recent study Spreng, MacKenzie and Olshavsky (1996) extended the disconfirmation of expectations theory to include desires by proposing a new model which integrates desires and expectations. Tse and Wilton (1988) suggested that, in addition to the influences from expected performance and subjective disconfirmation, "perceived performance exerts direct influence on satisfaction (p. 204)."

Although the dominant paradigm guiding recent research on satisfaction has been disconfirmation of expectations, some consumer researchers have been challenging and expanding

the disconfirmation paradigm and suggesting that many other determinants also affect satisfaction (Erevelles & Leavitt, 1992; Woodruff, *et al.*, 1983). Attributions about causes of performance (Folkes, 1984; Weiner, 1980) and the effects of amount of involvement (Zaichkowsky, 1985), for example, have also been shown to affect or modify satisfaction.

Most studies to date on consumer satisfaction have been limited to exploration of dimensions of consumer behavior regarding goods and services and have only occasionally focused on elements of the physical environment (Cadotte & Turgeon, 1988; Pate, 1993; Bitner, 1990; Donovan & Rossiter, 1982). Bitner (1990) stated that "Research in consumer behavior has tended to focus on the role of core attributes of products/services in determining customer responses. Relatively little empirical work has been done to examine the role of what might be considered peripheral variables, such as the physical surroundings in which the product/service is consumed. . ." (p.79-80). Although Bitner referred to physical surroundings as "peripheral," in this study the researchers considered them to be the focus.

A study by Caughey, Nafis and Francis (1995) examined the effect of involvement upon satisfaction with a university dining facility. Of four items measuring various aspects of involvement, only whether or not the subjects studied in the facility was useful in explaining satisfaction. Studies of retail environments have suggested that pleasant store atmosphere may influence purchasing behavior, or purchase intentions. Donovan and Rossiter (1982) suggested that store atmosphere can affect the pleasure and arousal states of consumers within the store. They contended that, "Whereas cognitive factors may largely account for store selection and for most of the planned purchases within the store, the emotional responses induced by the environment within the store are primary determinants of the extent to which the individual spends beyond his or her original expectations (p.54)."

Satisfaction with the built environment on university campuses is an important research area because of the recent tendency by some observers to view the university in a business/corporate model and thus to regard students as consumers of education (Levine, 1997). Those in the hospitality industry have long been aware that the environment in which they offer their goods and services has an important effect upon the attraction of customers and upon their purchasing behavior. Because of this awareness, they usually operate within a remodeling cycle of 3 to 5 years rather than the much longer cycle in the professions such as law and higher education. In recent years some professionals too have begun to refurbish their offices and facilities because it has become evident to them that the appearance of care about their near environment increases client confidence in their services. In the highly competitive field of the design of long-term healthcare facilities, owners of these facilities have emphasized a homey, comfortable ambience because they have observed that the families of the elderly potential clients are attracted to such amenities.

The evaluation of service encounters related to consumer satisfaction has been undertaken in several studies such as those of Bitner (1990, 1992) and Pate (1993). Pate (1993) noted that the recent paradigm of student satisfaction with universities views education as a service purchased by the consumer and that satisfaction in education "can be defined generally as satisfaction with the service(s) provided by the university or college" (p. 103). Bitner's later study (1992) explored factors which influence customer evaluations of service encounters including employee responses to complaints and the effects of the physical surroundings on customer satisfaction with the service encounter. Bitner (1990) suggested that research should be done to explore the measurement of the likely impact of changes in the physical design of a facility and how frequently the physical environment should be upgraded and what the benefits of remodeling would be (p. 80).

The present objective was to examine the effect of user expectations on their satisfaction with an interior space. The preliminary proposition explored was that the subjects' expectations about the dining facility would be related to their satisfaction with it.

METHOD

The focus of the study was a recently remodeled student-owned cafeteria in the student union building of a west coast university. The original facility, built in 1959, featured a vaulted, beamed ceiling, low levels of artificial lighting, an inadequate ventilation system which dispersed deep-fat fryer odors throughout the dining facility, a congested scramble area where users selected food, traffic bottlenecks at the cashier's stands, limited natural light, virtually no seating availability at peak dining hours, broken furniture, and stained carpet.

In 1994 the dining facility was closed for a \$2.5 million refurbishing, designed by a regional architecture firm with the aid of input from users of the facility. The seating capacity was increased from 300 to 500 customers, the amount of natural and artificial lighting was increased, more small tables were used, the scramble area was better planned and included a vaulted skylight, the ventilation system was upgraded, new furniture was purchased, and the food service format was changed to include national fast food franchises.

The dependent variable in the study was user satisfaction with the facility (Appendix 1, item #10). Overall satisfaction with the facility was measured on a 5-point scale (1 = not satisfied; 5 = very satisfied). It should be noted that the focus of this study was satisfaction with the built environment rather than satisfaction with food quality which some may consider to be the dominant influence in satisfaction with a dining facility. For our analysis, questions on the survey which referred to the physical appearance of the dining facility were used (items # 1, 1a, 8, and most of the items in #9). However, item #10 which asked subjects to indicate their overall satisfaction with the facility did not instruct subjects necessarily to consider only the physical environment. Therefore, it is possible that subjects' responses included their assessments of food quality, service quality, and so forth, as well as their assessments of the physical environment.

The independent variable was users' expectations. Expectations were assumed to vary depending on whether or not a user had seen the facility prior to the remodel. Expectations were operationalized as whether or not users had seen

the facility prior to the remodel. A questionnaire was developed (See Appendix 1) in which subjects were asked whether or not they had seen the facility prior to the recent remodeling (item #1).

Because expectations (i.e., whether or not one had seen the original facility) occurred naturally rather than as a result of experimental manipulation and because random assignment could not be employed, the design is considered pre-experimental. Because the facility was in such poor condition prior to remodeling, it was believed that users who had seen the facility before the remodel would have higher expectations regarding the new facility than would those who had not seen it. It was therefore expected that those who had seen the original facility would report higher satisfaction with the newly remodeled facility than those who had not seen it. The major threat to internal validity associated with this type of design is selection (Campbell & Stanley, 1963). That is, the researcher cannot assess the extent to which the subjects would have varied in their satisfaction with the facility regardless of their expectations. However, the value of independent variables in this realistic field setting was believed to outweigh this disadvantage. The effect of expectations on user satisfaction was analyzed using analysis of covariance and correlation coefficients.

Finally, subjects who indicated that they had seen the facility prior to the remodeling were asked to indicate to what extent they expected the newly remodeled one to look different (1 = look the same; 5 = look totally different) (Appendix 1, item #1a). They were also asked if they thought it actually looked better (1 = doesn't look better; 5 = looks better) (item #8). A correlation coefficient was calculated to assess the relationship between these two variables.

Demographic information was also collected. Subjects were 132 university students selected by a modified snowball sampling technique. Students in a sophomore-level interior design course filled out one questionnaire themselves, then asked two other students to fill out questionnaires. All were current users of the student dining facility. Because the instrument was based on recall, students were asked not to administer the instrument while in the facility. As a result of the sampling procedure, 33% of the sample were either interior design or housing design majors; the remaining 67% were

other majors. To check for any possible effects of major, preliminary analyses were performed. Students' majors were recoded into two categories: a) interiors/housing and b) all other majors. No significant differences between the two categories of majors were found for either the extent to which they expected the newly remodeled facility to look different (Chi square = 0.58, $df=3$, $p=0.90$) or the extent to which they thought it actually looked better (Chi square = 6.18, $df=4$, $p=.19$). However, a significant difference between majors was found for overall satisfaction with the remodeled facility (Chi square = 11.30, $df=4$, $p<.05$). Therefore, major was taken into account in the subsequent analysis of overall satisfaction.

Forty-nine percent of the subjects had seen the facility before the remodel. Women comprised the majority of users (62%). Users' ages ranged from 18 to 42 years, with a mean of 21.5 years. Academic majors represented eight of the ten undergraduate colleges. Ninety percent of the subjects were American citizens. Seventy-four percent of the subjects who were American citizens were Caucasian, 21% Asian, and the remainder comprised several ethnic/racial minorities. Class ranks were 13% freshmen, 23% sophomores, 33% juniors, 27% seniors; and the remaining 4% were graduate and unclassified students.

RESULTS

Subjects were asked to report their level of overall satisfaction on a scale of 1-5 (Appendix 1, item #10). As the means for each of the two groups were above 3.5, it should be noted that neither group reported general dissatisfaction with the newly remodeled dining facility. To examine the proposition that users' expectations would have an effect on satisfaction with the dining facility, an analysis of covariance was conducted. The dependent variable was overall satisfaction. The independent variable (main effect), expectations, had been operationalized as whether or not users had seen the facility prior to the time it was closed for remodeling. Major was used as a covariate. Results are displayed in Table 1. Expectations had a significant effect on overall satisfaction with the facility ($F=4.66$ $df=1, 122$; $p<.05$) with prior differences due to the effect of major ($F=12.52$

$p < .001$) removed.

Table 1
Analysis of covariance for Satisfaction
by Expectation

Source	Sums of Squares	d.f.	Mean Square	F	p
Covariate					
Major	9.77	1	9.77	12.52	.0006
Main Effect					
Expectations	3.63	1	3.63	4.66	.0329
Residual	95.21	122	0.78		
Total	110.27	124			

The overall satisfaction mean for the group of users who had seen the dining facility prior to remodeling was 3.59 ($n=61$). For those who had not seen the facility prior to remodeling, the satisfaction mean was 3.94 ($n=64$). That is, those who had not seen the dining facility prior to the time when it was closed for remodeling reported significantly higher overall satisfaction with the remodeled facility than did those who had seen the unremodeled facility.

To look more specifically at the direct relationship between expectations and satisfaction, the correlation between the extent to which users expected that the remodeled facility would look different and the extent to which users considered the remodeled facility actually to look different was calculated. For this analysis, only those users who reported that they had seen the facility prior to the remodeling were included. A significant positive correlation was found between expectations and satisfaction ($r = .25$, $p < .05$). As the extent to which it was expected that the remodeled facility would look better increased, the extent to which the remodeled facility was assessed as actually looking better also increased.

DISCUSSION

As suggested by disconfirmation theory, in the present study expectations had an effect on satisfaction.

According to Woodruff, Cadotte, & Jenkins (1983) and others, satisfaction occurs based on a comparison of that which is expected with that

which is received. In the current study, users who had seen the facility in its previous state were considered to have had more information than those who had seen only the newly-remodeled facility. That is, their previous exposure and concomitant greater amount of information logically should have influenced their satisfaction with the newly remodeled facility, as was found to be the case. Specifically, users who had seen the facility before remodeling had been expected to be more satisfied with the newly remodeled facility than those who had not. However, the reverse was found. Users who had seen the facility prior to the remodeling were less satisfied with the newly remodeled facility than were those who had not.

It was previously noted that the original facility was in poor condition prior to the remodeling. It may be that users with previous exposure held negative impressions of the facility which led to the formation of negative expectations of the new facility which, in turn, resulted in lower overall satisfaction compared to users without previous exposure. Perhaps negative prior experience generated lower expectations which resulted in lower satisfaction. Some support for this proposition may be found in the Hunts' (1988) study of consumer grudgeholding in which it was reported that consumers did form and retain grudges as the result of negative experiences. These grudges were found on a widespread basis among consumers and were held tenaciously. It would be logical that any subsequent product/service expectations would be correspondingly lower than might otherwise be the case. Perhaps such was the case in the present study.

Another possible explanation for the current findings may be based on the characteristics of the two groups of users. That is, perhaps there were some underlying initial differences between the two groups that may account for the difference in their overall satisfaction.

This raises the question of how expectations about facilities are formed, and whether this formation is different from expectation formation regarding goods and services.

A suggestion for future research would be to conduct a laboratory experiment to study both heightened and lowered expectations to allow

manipulation of the dependent variables. Perhaps heightened expectations affect satisfaction but lowered expectations do not. This proposition is worthy of further investigation. Or, in future studies expectations of potential users could be measured before they see a facility because once they have seen it it may be difficult for them to separate their previous expectations from their actual impressions. A different measure of expectations such as to what extent users expect to be completely satisfied with a newly remodeled facility might be a better measure of expectations than was used in the current study.

The basis for another study also could be the determination of the relative importance of food quality and quality of the built environment in determining satisfaction with dining facilities.

A more general issue is that perhaps interiors and facilities do not generate expectations in the same manner in which products and services do. Studies could explore what aspects of interiors, if any, contribute to overall satisfaction and whether expectations about other aspects of a facility such as food or service quality contribute to satisfaction with interior environments.

REFERENCES

- Bitner, M. J. (1990), "Evaluating Service Encounters: The Effects of Physical Surroundings and Employee Responses," *Journal of Marketing*, 54, 69-82.
- Bitner, M. J. (1992), "Servicescapes: The Impact of Physical Surroundings on Customers and Employees," *Journal of Marketing*, 56, 57-71.
- Cadotte, E. R. and N. Turgeon (1988), "Key Factors in Guest Satisfaction," *Cornell Hotel and Restaurant Administration Quarterly*, 45-51.
- Campbell, D. T. and J. C. Stanley (1963), *Experimental and Quasi-Experimental Designs for Research*, Chicago, IL: Rand McNally.
- Caughey, C. C., D. A. Nafis and S. K. Francis (1995), "Level of Involvement in Relation to Satisfaction with a Student-Owned Dining Facility," *Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior*, 8, 234-238.
- Donovan, R. J. and J. R. Rossiter (1982), "Store Atmosphere: An Environmental Psychology Approach," *Journal of Retailing*, 58, (1), 34-57.
- Erevelles, S. and C. Leavitt (1992), "A Comparison of Current Models of Consumer Satisfaction/Dissatisfaction," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 5, 104-114.
- Folkes, V. S. (1984), "Consumer Reactions to Product Failure: An Attributional Approach," *Journal of Consumer Research*, 10, 398-409.
- Hunt, H. K., H. D. Hunt and T. C. Hunt (1988), "Consumer grudgeholding," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 1, 116-118.
- Levine, A. (1997, January 31), "Higher Education's New Status as a Mature Industry," *Chronicle of Higher Education*, 48.
- Pate, W. S. (1993), "Consumer Satisfaction, Determinants, and Post-Purchase Actions in Higher Education: A Model to Guide Academic Managers," *College and University*, 68, 100-107.
- Spreng, R. A., S. B. MacKenzie and R. W. Olshavsky (1996), "A Reexamination of the Determinants of Consumer Satisfaction," *Journal of Marketing*, 60, 15-18.
- Tse, D. K. and P. C. Wilton (1988), "Models of Consumer Satisfaction Formation: An Extension," *Journal of Marketing Research*, 25, 204-212.
- Weiner, B. (1980), *Human Motivation*, New York: Holt, Rinehart & Winston.
- Woodruff, R. B., E. R. Cadotte and R. L. Jenkins (1983), "Modeling Consumer Satisfaction Processes Using Experience-Based Norms," *Journal of Marketing Research*, 20, 296-304.
- Zaichowsky, J. L. (1985), "Measuring the Involvement Construct," *Journal of Consumer Research*, 12, 341-352.

Appendix 1

Questionnaire

1. Did you see the MU Commons before June, 1994, when it was closed for remodeling? (Circle one.)
 - a. no
 - b. yes
 - 1a. If yes, to what extent did you expect that the remodeled MU Commons would look better: (Circle one.)

Expected it to look the same	Expected it to look totally different
1 2 3 4 5	
2. Have you been to the MU Commons since September, 1995? (Circle one.) If no, skip #3-10, and turn to the last page.
 - a. no
 - b. yes
3. On the average, how many hours per week do you spend in the MU Commons: (Circle one.)
 - a. less than one hour a week
 - b. 1-3 hours a week
 - c. 4-6 hours a week
 - d. more than 6 hours a week

BUSINESS LINK: A METHOD FOR EXAMINING CUSTOMER SATISFACTION, DISSATISFACTION & COMPLAINING BEHAVIOUR

Susan J. Priest, University of Cambridge

ABSTRACT

This paper develops the methods for assessing customer satisfaction, dissatisfaction and complaining behaviour (CS/D & CB) for application to the evaluation of the British Business Link initiative. Business Link is an initiative to provide advice and consultancy to small business clients. Methods of using CS/D have been well developed since the 1970s as means to evaluate new product and service developments. Their application to public policy is less common. In this paper methods of CS/D & CB evaluation are developed for Business Link contrasting this approach to the Government's own assessments. The paper argues that assessing Business Link by measuring CS/D & CB in the manner proposed in this paper is the most rigorous approach for policy planning and evaluation purposes.

CUSTOMER SATISFACTION & DISSATISFACTION

The importance of assessing customer satisfaction and dissatisfaction (CS/D) has been documented extensively since 1970, when the United States Department of Agriculture's Index of Consumer Satisfaction (Pfaff, 1972) reported on CS/D information to policy makers (Churchill & Surprenant, 1982). Much of the early work on the antecedents of CS/D was undertaken by Cardozo (1964), Olshavsky & Miller (1972) and Anderson, (1973). Since that time numerous theoretical structures have been proposed and empirical studies undertaken with the aim of improving the information available for research, public policy, consumers, and markets (Hunt, 1976).

In 1976 Greyser noted that there was inadequate conceptualisation and measurement of CS/D. He proposed CS/D research must be grounded in the use of CS/D: (1) as a social indicator - an index providing a longitudinal measure for noting changes in CS/D over time; (2) as central to, and a goal of, an economic system; (3) as a guide to policy, planning and evaluation, to give policy makers conceptual and measurement support for their decision making; (4) as a guide in

specific regulatory matters; (5) to support public policy making decisions in consumer legislation; (6) as a central precept of the marketing concept to support the marketing strategy of a business; (7) in consumer activities allowing activists to focus attention on those aspects which show the lowest level of satisfaction; and (8) as a basis for developing theories of consumer behaviour.

To develop these various bases for application a wealth of approaches to CS/D has been developed. One widely accepted conceptualisation of CS is a measure of whether it meets or exceeds customer expectations (Anderson, 1973; Oliver, 1980, 1981). This has provided impetus for a number of studies that use a disconfirmation paradigm as a view of the process by which customers develop feelings of satisfaction or dissatisfaction. As Bateson & Wirtz (1991, p. 4) note "...in this paradigm, consumers evaluate consumption experiences and make satisfaction decisions by comparing perceived performance with some pre-consumption standard. The level of satisfaction is related to the size and direction of disconfirmation experienced".

Although most empirical research has modelled and measured CS/D in a commercial setting for consumer goods and services at the end of the supply chain (Spreng, MacKenzie & Olshavsky, 1996), its application to public policy has been less well developed. The application for accurate and reliable assessment of CS/D for the supply of business services delivered through organisations offering business support should be straightforward. However there has been no previous attempt to apply a comprehensive assessment of CS/D to this type of problem. It is this gap that the paper addresses. The paper first reviews the underlying concepts and the associated debates surrounding the public policy to develop services for small and medium sized firms in Britain, branded as Business Link. It then describes the main current evaluations being applied to this initiative, none of which make comprehensive use of CS/D & CB. The paper then develops a means of measuring CS/D & CB for the case of Business Link, arguing that this is the most comprehensive and rigorous assessment for

policy planning and evaluation purposes.

ENTERPRISE SUPPORT FOR SMALL & MEDIUM-SIZED FIRMS IN BRITAIN

Business Link is the latest in a line of initiatives by which the government has sought to help small and medium-sized firms develop in Britain. These initiatives began with the publication of the Bolton Committee report in 1971 which considered the role of small firms in the national economy, the facilities available to them and the problems confronting them. The Committee examined in particular the profitability of small firms, the availability of finance, and the special functions of small firms, for example as innovators and specialist suppliers.

Since that time numerous policy initiatives have been launched to help small firms overcome what the Bolton Committee noted as "a number of inequitable and unnecessary disabilities, mostly imposed by government which amount to discrimination against them" (Bolton, 1971 p. 91). The first of these initiatives was the Small Firms Counselling Service. This was primarily a telephone-based signposting and referral service which started in 1979 and ran until 1990, after which the budget was absorbed into the newly developed Training & Enterprise Councils in England and Wales and the Local Enterprise Companies in Scotland. During this eleven-year period a multitude of other initiatives were launched to stimulate the small firms sector. The net result was duplication of activity from a confusing range of agencies all with different primary objectives. This situation continued until there was a major review of government policy in 1992 which sought to increase the involvement of the private sector *with* government in support of enterprise (Titchener, 1996). The fundamental conclusion of this review was that existing support schemes lacked customer orientation with the result that services were too narrowly defined and more supplier- than customer-driven. There was also a low level of awareness by small and medium-sized firms of the support services available: many firms that could benefit from business support services were unaware of their existence.

The Business Link network was borne from

this review. They were launched with the key aim to "simplify the confusing array of support services"... and provide "a single point of access to the highest calibre support services" (Michael Heseltine, 1992, p. 415; then President of the Board of Trade). The launch of this new strategy by the government in July 1992 was an attempt to "improve the competitiveness of companies and provide a springboard for the development of local and national economies" (Heseltine, 1992, p. 415). Since the launch Business Links have subsequently developed into a one-stop-shop, encouraging firms to access external support through a single channel, formally branded as Business Link in England. Variations on the concept have emerged as "Business Shop" in Scotland and "Business Connect" in Wales.

Objectives of Business Link

The Business Link network was developed with a number of key objectives in mind (DTI, 1992). The first sought to provide greater coherence between the existing main suppliers of enterprise services particularly the Training & Enterprise Councils, Chambers of Commerce, Enterprise Agencies and local government. These bodies were drawn together as 'partners' in supplying Business Link services. The second objective sought to improve the professionalism and quality standards of the support services available across the country by combining government resources with local inputs from partner organisations of the Business Link and from private sector sponsorship. Third, a core focus for the Business Link is the new general business adviser, the Personal Business Adviser, who provides the one-stop-shop or diagnostic service that allows Business Link to act primarily without further referral. Fourth, the initiative sought to create a network of outlets to have a physical presence across the country. These objectives with an overt shift towards a 'partnership' approach to enterprise support have now meant that many of the existing local business support organisations have become the main partners of Business Link, a number of which are co-located and operate within or as part of the Business Link companies. Due to closer working relations many partners have contractually agreed

to link funding and staff into the single Business Link brand, whereas others have dual or joint branded services with the Business Link.

Throughout the launch and establishment of the network a number of issues have been raised by commentators about the initiative. The overriding concern has been that Business Link has been subject to a high level of government central planning and direction which has questioned whether the policy design and management can be sufficiently customer focused. An early assessment of the design by Bennett (1995) warns of four main areas of potential failure. First, the misdesign of incentives and management where "effective contracts between principle and agent require a number of simple but key principles to be satisfied" (p.36). He demonstrates that excessive control and prescription are likely to lead the Business Link to violate the basic principles for the management of an efficient and effective organisation. Second, the market demand for some of the new services is not clear. For example, it is not clear that the central and new service of the Personal Business Adviser is needed or sustainable. Third, there are high management costs and a lack of critical mass. Dividing government resources among approximately 90 companies operating from over 200 outlets in England is likely to lead to "penny parcels" (p. 34) which increases staff and premises costs and diffuse management. Fourth, Business Link has the potential to stimulate adverse selection and moral hazard, i.e. they are designed in such a way that they have the potential to target and support firms who may be those least viable, stimulating firms to take unnecessary risk. Finally Bennett (1995) highlights how the core new Personal Adviser role operates on account manager principles where they are encouraged to act selectively, in that they have targets of assisting potential growth firms, while also coming under increasing pressure to secure client fees. Overall this is distortionary to the market by giving specific aid to some firms rather than others, the net result may be to stimulate a dependent business community which ultimately may harm the growth and competitiveness of the economy, in effect achieving quite the reverse of Heseltine's original objective for Business Link.

In part response to such criticisms, and as part of its wider policy evaluation, a number of

investigations have taken place. The government is constantly monitoring the performance of Business Link and has commissioned a number of reports including monitoring studies (Carma, 1995); evaluation studies (KPMG, 1994; Ernst & Young, 1996; ISG, 19/96); impact studies (DTI, 1996b, 1997c); service measurement studies (Mall Research Services, 1994); best practice assessments (Browne, 1995); mystery enquirer research (BPRI, 1997); training assessments (Mall Research Services, 1998) and studies on the awareness and use of the Business Link network (MORI, 1996). In addition the British Parliament's House of Commons Trade and Industry Select Committee has investigated Business Link with particular reference to the quality of service, focus of services for start-up and micro-businesses, funding and long term viability, and relationships with other organisations (HC 302-I, 1996). Perhaps the most pertinent comment the Committee made is that customer satisfaction surveys used by Business Links are inadequate and unreliable indicators of performance as they do not seek to establish customer expectations in relation to the satisfaction construct (HC 1996, p. 49, para. 144). In addition to this criticism, they made a total of thirty-two recommendations that covered a number of structural aspects of government contracting and relations with Business Link and partner organisations. Despite a change of government in May 1997, no fundamental change in the design features of Business Link has yet been implemented. A revised five-year vision statement for Business Links (DTI, 1997a) from the new government was disappointing in that it pledged only limited modifications and hence it is likely that the debate surrounding contract reform, the design of the services, the viability of Business Links, and their potential to be customer focused will continue.

As part of the underlying rationale of Business Link is to provide customer focused and highly responsive services to help support the development needs of business, it is surprising that these assessments do not use theories of consumer behaviour to provide a fuller evaluation of the policy. A distinct lack of attention to this as a way of explaining customer responses and improving the services offered through the

initiative represents a major flaw. Hence this paper focuses on how to evaluate Business Link using the methods of CS/D & CB. It argues that as the assessment of the value and impact of the support can only properly rest with the business itself (Bennett, 1996), measuring CS/D & CB in the manner proposed here is the most suitable and rigorous approach for this policy evaluation. It also identifies that significant improvements can be made by integrating client feedback from the assessment into the local and national marketing strategies for the initiative.

MEASURING CS/D & CB FOR BUSINESS LINKS

As mentioned earlier, the Department of Trade & Industry (DTI), as the main contracting and funding partner, has commissioned a number of studies that evaluate different aspects of the initiative. Of particular interest is that they recommend the use of CS/D measurement as a tool by which the Business Link partnerships can assess their own performance (ISG 13/96). In addition information on CS/D is required for the quarterly monitoring statistics supplied via the Training & Enterprise Councils to the DTI (DTI, 1997b), and as part of the accreditation framework for the Business Link brand. For example one target specified in the funding offer letter is "at least 85% of customers satisfied or very satisfied". Also as part of the ISO 9001 criteria Business Link has to demonstrate that the survey work it carries out is robust. To encourage this assessment the DTI issues guidelines on interpreting ISO 9000 (ADEPT, 1995), yet they provide wholly inadequate information on how to undertake quantitative and qualitative research on customer satisfaction (ISG 13/96). This limited support gives a very brief overview of sampling, non-response bias, piloting, the different types of surveys, questionnaire design, and organising focus or discussion groups. The DTI does not consider a number of *essential* and underlying themes necessary for a comprehensive and rigorous assessment (1) how Business Link identifies and establish customer expectations, or the comparison standards customers use in the S/D judgement; (2) how Business Link can measure the effect of, size and direction of disconfirmation to

explain the level of the satisfaction response; (3) the value to be gained from in-depth understanding of consumer complaining behaviour; and (4) how Business Link can integrate client feedback into its marketing strategy to improve service quality. As a direct consequence across the network much of the current practice is inadequate, unreliable and is highly variable in its quality and content.

The remainder of this paper highlights the differences in approach of that promoted by the DTI and that possible through a CS/D & CB approach. The discussion conveys four main aspects. The first section proposes a method to examine customer expectations and or other comparison standards used by consumers in making the satisfaction judgement. It argues that this diagnostic element is vital in the effective management of customer relations. The second section evaluates a number of key client focused quality dimensions that are generic to this type of service interaction, many of which are acknowledged as important to the success of service-based organisations. It also measures CS/D for a range of contextual issues about the client's business that arise from the design of the policy. The third aspect proposes a series of specific impact dimensions for evaluation that derive from aspects of the underlying rationale of Business Link. These seek to complement the CS/D measures to provide a more rigorous assessment of whether the policy is achieving its objectives. The final section of the paper proposes how to examine the outcomes of Business Link performance by detailing a number of measures.

Identifying Customer Expectations or other Comparison Standards

Although a range of CS/D measures will rate the advice satisfaction and will give an indication of the adviser's performance, for policy evaluation purposes it may be limiting to consider only the importance of the percentage of satisfied or dissatisfied customers (Woodruff & Gardial, 1996). As the House of Commons Select Committee noted (HC 302-I, 1996), customer satisfaction surveys of Business Link need also to establish customer expectations in relation to the satisfaction construct, an opinion the DTI does not endorse. However numerous theoretical and

empirical studies have proposed that performance can be compared to standards *other* than expectations in the process leading to the satisfaction judgement. As a result the expectations disconfirmation model is not the only model used to assess CS (see e.g. the work of Oliver, 1980; Woodruff, Cadotte & Jenkins, 1983). Further to this, Jayanti & Jackson (1991, p. 603) suggest that "when performance judgements tend to be subjective (as in the assessment of most business services due to intangibility) expectations may play a more minor role in the formation of satisfaction", or as Patterson, Johnson & Spreng (1997, p. 6) suggest, "satisfaction in services may be a function of performance alone".

While many of these authors have tested for the effect of different standards for particular products or services in a range of settings, a number of comparison standard categories can be identified. For Business Link purposes it is important to establish which of these standards is most relevant to their customers. This will help service delivery personnel and marketing executives understand how customers develop a perception of the service and what they anticipate will happen during the interaction. Although there is a range of comparison standards some are more relevant in this context than others, for example expected performance which is derived from expectancy theory is the most commonly used standard and has the strongest empirical support (Bateson & Wirtz, 1991). It may reflect what the customer feels performance of the Business Link advisers "probably will be" (Miller, 1976), or what the consumer predicts performance will be (Woodruff, Cadotte & Jenkins, 1983). From a Business Link viewpoint this may be one of the more easier standards to manage, control and ultimately satisfy as more recent work by Spreng, MacKenzie & Olshavsky (1996) has conceptualised predictive expectations as being easier to define and identify. From a business advisers viewpoint these may more easily be satisfied by a similar level of performance because they are potentially more explicit or overt.

Another comparison standard is ideal performance. While this is derived from the models of the ideal point of consumer preference and choice (Holbrook, 1984), it represents what

the Business Link consumer feels performance "can be". As it also represents an optimistic view of the potential interaction, it is proposed here that this may be more difficult for the Business Link adviser to manage and subsequently satisfy as it more closely reflects an overall 'desire'. Again as Spreng, MacKenzie & Olshavsky (1996) suggest these may consist of more nebulous concepts that are also potentially more discrete. In a business development context they may represent an unrealistic perception of performance as there are complex and tenuous relations between the provision of external information or advice, firm management and business growth. It is proposed here that Business Link customers making reference to an ideal, e.g. using the services to stimulate business growth, may be more difficult for the adviser to satisfy in contrast to those making comparison to expectations that are more clearly defined.

A third comparison standard with potential application is normative performance. While this is derived from equity theory (Adams, 1963), it reflects the level of performance a consumer "ought to receive" or "deserves" given a set of costs (Miller, 1976; Woodruff, Cadotte & Jenkins, 1983) and is determined by the consumer's evaluation of the rewards and costs or investments and costs. In this context Business Link clients may evaluate the cost of the interaction with the Business Link adviser either in pure monetary terms by fee-paying for the support, or through time spent with the adviser (if no fees are charged) compared with the output they experience. Value for money and value for time are commonly used slogans to give a fairly basic measure.

The final standards which may be relevant to Business Link are those based on experience. While these are formed through prior experiences (LaTour & Peat, 1979), word of mouth endorsements, criticisms, and / or marketing effort (Woodruff, Cadotte & Jenkins, 1983), it may also reflect desired performance in meeting consumers' needs and wants while being constrained by the view of performance that consumers feel is possible based on the breadth of consumption experiences (Bateson & Wirtz, 1991). As many Business Link customers will have used other external sources for business information and support, experience may contribute to the

development of these standards which then play a significant role in the evaluation of Business Link performance.

Hence the *diagnostic value* of this CS/D research is dependent on the examination how Business Link customers make comparison using these standards. A CS/D evaluation of Business Link will differ significantly from that undertaken by the DTI by focusing first on identifying if and which comparison standards are being used by customers, or whether satisfaction with the service a function of performance alone. It also examines which of the comparison standards is being used, e.g. are they broader than just expectations of the service and which is the most dominant overall. The proposed evaluation assesses what customers anticipate or hope would happen as a result of using the Business Link service to establish their predictive expectations or desires from a government-backed initiative. It also examines whether customers making reference to the ideal are more demanding or difficult to satisfy. The implication is whether greater attempts should be made to create realistic and predictive expectations that are sufficiently demanding of the adviser by the average level of performance. It also secures a measure of performance by examining if the service failed, met, or exceeded those expectations. To provide a fuller analysis the customers' previous or starting expectations of performance is measured and integrated into the analysis to examine whether they experience significant variation in the size and direction of disconfirmation. Conceptualising this as a position on a disconfirmation continuum will provide a more accurate assessment of the effect of actual performance. Finally this approach examines the perceived value from the interaction with the Business Link, and for experienced users of external support services, how their assessment of Business Link performance compares to that of other providers.

A major criticism of the DTI and Business Link is that if they do not know if and which comparison standards are being used, they cannot manage customer relations effectively as empirical research has shown that each comparison standard is affected by a number of inputs, e.g. by the average product or service performance (Miller, 1976), advertising effects (Olson & Dover, 1979),

word-of-mouth communications and personal experience (Woodruff, Cadotte & Jenkins, 1983). Assessment of CS/D has greater value than that undertaken by the DTI as once the comparison standards have been identified, information relating to the inputs can be fed-back into the central and local marketing strategy of the initiative and for individual services. Adoption of a CS/D approach will therefore allow the DTI centrally and Business Links locally to manage customer relations in a more effective way, giving more accurate and reliable indicators of actual performance, and contribute more effectively to policy planning and evaluation.

The paper does not evaluate the psychological constructs of comparison standards, or develop any of the models of consumer satisfaction (for a review see e.g. the work of Erevelles & Leavitt, 1992). It is proposed here that knowledge of the psychological constructs is not necessary for the measurement and tracking of customer satisfaction over time, which can be accomplished by policy executives in Business Link without an in-depth understanding of each construct of the comparison standards.

Measuring Customer Satisfaction & Dissatisfaction with the Business Link Initiative

As Business Link is a provider of advisory services to business the success of the initiative is dependent largely on the performance of its advisers. Therefore part of any policy evaluation needs to concentrate on assessing the performance elements that contribute to advice satisfaction. To a limited degree the DTI has recognised this and issues to each accredited Business Link numeric requirements for six client-focused quality dimensions. The minimum requirements applicable across the network are shown in Table 1. While the DTI elaborates on each dimension by recommending for particular service types the minimum requirements and data sources, they encourage the use of closed (yes/no) relies thereby limiting the responses. The DTI assessment leaves no room for the reporting of mediocre or high variation performance. In contrast this paper argues that a multi-point (five-point) measurement scale of CS/D is essential to allow proper evaluation of advice satisfaction. This allows

assessment of a number of generic and key client focused performance elements including (1) the overall performance of the adviser; (2) the overall professionalism of the adviser; (3) the responsiveness of the adviser; (4) the charging strategy for the service; and (5) value for money. Many of these elements have been shown to be of great importance to the success of service-based organisations (Tucci & Talaga, 1997). Similarly the DTI does not recommend or evaluate measures of CS/D for contextual issues about the client's business that arise from the initiative. This suggests the need for three further dimensions of assessment (6) the advisers accurate analysis of the business support needs; (7) the advisers understanding of the business; and (8) the advisers understanding of the market(s) within which the firm operates. These are potentially even more important indicators of the adviser's performance. Accurate analysis of the firm's operating environment including a full appreciation of the strengths, weaknesses, opportunities and threats facing the client's business will ultimately affect whether the firm is given or referred to the correct source of support. This in turn will influence the customer's evaluation of performance and it seems to be an aspect the DTI has yet to recognise.

Assessing Specific Dimensions of Business Link: Impact Measures

As mentioned earlier, Woodruff & Gardial (1996) note that for policy planning and evaluation purposes it may be limiting to consider only the importance of the proportion of satisfied or percentage dissatisfied customers. CS studies commissioned and published by the DTI and Business Link continually report a high proportion of satisfied customers, which may be explained by the use of inadequate measurement scales. However in order to apply stronger rigour to this evaluation advice satisfaction measures are supported with impact data, as a satisfaction rating alone does not address whether the interaction with Business Link has had a major impact and what the effect may be to help improve the firm's competitive position: a key original aim of Business Link.

A more rigorous assessment should use both CS/D and impact measures on the same sample

group. It should assess whether the Business Link advisers are performing satisfactorily while also measuring the impact of the support on the performance of the firm. The evaluation proposed relies on the clients' qualitative assessment of the impact of the service on their business in an approach similar to that reported in Chrisman & Katrysten (1995). It is different from the assessment proposed by the interdepartmental group of the DTI (1997c) which recommends the use of control and comparison groups, tracking and comparing a number of key and secondary performance indicators at a distance from the firm. Consequently the assessment proposed here does not have the difficult (or perhaps impossible) task of matching and controlling for innumerable significant differences between comparison firms.

As 'impact' can be interpreted in a number of ways it can also be measured using a variety of methods. The DTI recommends that Business Link uses two indicative measures of impact (see Table 1). This proposed evaluation uses a more rigorous approach than just "50% of clients reporting the service benefiting their business", and "85% of action to be undertaken by Business Link, where agreed with client, achieved". Instead it should include the client's assessment of (a) customer views of the most important aspect(s) of the service; (b) general impact measures using descriptive terms such as the overall usefulness, descriptive impact of the service now and in the medium term, and added value; and (c) specific impact measures on the turnover, employment, profitability, and productivity of the firm now and in the medium term. These impact measures will support the satisfaction assessment to give an overall evaluation of the policy, and unlike the DTI assessment, this evaluation will identify those areas of *greatest value* to customers. Feedback from this type of assessment should help Business Link realign its services to meet customer requirements, and provide evidence of whether it is fulfilling some of their original objectives. Subsequent impact assessment should look beyond the obvious measures identified from the underlying rationale of the initiative or the remit of the individual services to integrate aspects nominated by clients. This will provide a more relevant assessment of impact grounded in the value given by clients.

Table 1
Numeric Requirements for Business Link Client Focused Quality Dimensions

Client Focused Quality Dimension	Indicative Measures	Service Types	DTI Minimum Requirements	Possible Data Sources
Responsiveness	1. Service delivered within deadlines agreed between Business Link and client;	All	1. 80% delivered within scale;	Client management system/surveys.
	2. Clients contacted within promised time scale after receipt of service.	All	2. 80% contacted within scale.	
Accuracy	3. Answers to enquiries accurate. The latest available updates of information in use.	Information	3. 80% accurate response.	Internal audits. Audits of information sources.
Appropriateness	4. Service provided appropriate to client needs;	All	4. 85% appropriate service;	Market research and client surveys.
	5. Diagnosis of client need correct;	All	5. 90% accurate diagnosis;	
	6. Referrals to other organisations appropriate.	Referrals	6. 80% referrals appropriate.	
Professionalism of Service	7. Staff knowledgeable and provide service in a professional manner.	All	7. 5% max. unresolved complaints.	Client surveys/records.
Long-term Relationship	8. Clients returning to Business Link for other services.	All	8. 80% clients would or do return for future services as first choice.	Internal client management system. Membership records.
Impact	9. Service benefited clients business;	All	9. 50% reported benefit;	Client surveys. Audit of outcomes against action plans.
	10. Action to be undertaken by Business Link, where agreed with client, achieved.	PBA/ Consultancy	10. 85% achievement of outcomes of action plan.	

Source: DTI (1996a) Business Link Accreditation Booklet, p. 52.

Measuring the Outcomes of Business Link Performance

Finally in terms of measuring the outcomes of performance, the DTI recommends just one numeric target, which is that "80% of clients would or do return to Business Link for future services as a first choice" (Table 1). This is very simplistic. Although research has shown that there are complex relations between satisfaction,

complaints, and repurchase intentions (Jacoby & Jaccard, 1981) many firms, and in this case the DTI, view repeat custom or the intention to repeat custom as a positive outcome of performance. This is not necessarily a measure appropriate for government supported services such as Business Link that seeks to fill market gaps. As Bennett (1995) highlights, the encouragement of the continued use of Business Link services has the potential to create a business community dependent

on subsidised services, can create unfair competition with existing private sector suppliers, and hence may have the capacity to harm the growth and competitiveness of the economy, the reverse of its objectives.

To examine outcomes in greater depth and to assess whether Business Link is stimulating a dependency on subsidised services, a fuller assessment requires outcome measures that include the client's intention to use the same service again. This will examine if sufficient value has been added to warrant continual or repeat use, and to give a measure of the evaluation of the rewards and costs as perceived by the customer. It also includes the client's intention to use other Business Link or partner services in the future, thereby examining whether poor, mediocre or high variation performance is affecting the customer's intention to use similar branded services. In addition outcomes of performance should also measure whether the service has improved the *competitiveness* of the firm as this links to the original objectives of Business Link. Using this criterion for evaluation overlaps with some of the impact measures introduced earlier and acts as a check on response accuracy.

A further outcome measure that at present the DTI seems to undervalue is information relating to the reactions of dissatisfied customers through CB. CB has been the subject of numerous studies, some of which have noted the benefits from encouraging CB, while others have suggested that dissatisfaction is a complex phenomenon. Although a number of studies have shown that dissatisfied customers often tend not to complain directly to the manufacturer, retailer, or service provider (Naumann & Giel, 1995), customer comments are vital to any customer-focused organisation--an explicit aim of Business Link. Consequently it is surprising that the DTI does not give any explicit or detailed guidance to Business Link on this matter.

The proposed assessment makes a comprehensive evaluation of CB. It does not just note the main or overriding themes of dissatisfaction which is documented as part of the ISO standard recommended by the DTI. It seeks to capture as much information on CB as possible, identifying the cause of dissatisfaction and customer action as a result of feeling dissatisfied.

If customers do not complain to Business Link or elsewhere, or take any corrective action, this evaluation examines the corresponding reasons. If clients complain to the Business Link it identifies the corrective action taken as well as the desired corrective action. Finally customers are asked to rate their satisfaction or dissatisfaction with the way their complaint was investigated, and rate the final outcome of their complaint. This is relevant to Business Link as research reported in Naumann & Giel (1995) has shown that a substantial amount (70-90%) of complaining customers will repeat use if they are satisfied with the way their complaint was handled, and the DTI issues targets for the development of long-term relationships with clients that include measures of repeat business (DTI 1996a).

DISCUSSION

This paper has reviewed a selection of work on the conceptualisation and measurement of CS/D since the 1970s. It starts from Greyser's (1976) proposal that CS/D measures can be used for policy planning and evaluation purposes to give policy makers conceptual and measurement support for their decision making. It also identifies the value for Business Link of integrating customer feedback into the national and local marketing strategies. It draws on the more recent work of Woodruff & Gardial (1996) to argue that CS/D measures are rigorous and suitable measures if combined with other indicators of performance that arise from the underpinning philosophy or objectives of the policy. The paper applies these approaches to the case of Business Link to propose an independent assessment method as an alternative to that being used by the DTI. It suggests a more rigorous approach which (1) examines customer expectations and other comparison standards used in the S/D judgement to establish if and which standards are being used by customers. It also examines how actual performance measures up to these reference points through integrating the level or size and direction of disconfirmation into the analysis; (2) measures CS/D for a range of generic and contextual client focused performance dimensions taken from the underlying rationale of the policy; (3) collects a range of descriptive and quantifiable impact data

using a number of methods; and finally (4) measures the outcomes for clients of a number of aspects of Business Link performance.

The paper also highlights the differences in approach of that adopted and recommended by the DTI for Business Link implementation and what a full assessment of CS/D & CB requires. It shows how the DTI promotion of satisfaction measurement has a number of severe shortcomings. For example, although the DTI issues guidelines on how to undertake evaluations of CS they do not address a number of essential themes argued as necessary for a comprehensive and rigorous assessment of the policy. The main criticisms of the DTI approach are that (1) the numeric performance targets are limiting in that they encourage closed responses so that mediocre or high variation performance is not reported; (2) they do not evaluate satisfaction with contextual issues about the client's business and operating environment arising from the initiative, which it is argued here will undoubtedly influence the customer's evaluation of performance; (3) they do not make any attempt to understand how customers make the satisfaction judgement. The DTI evaluations do not have any diagnostic value; they do not assess if clients are using comparison standards, and once identified, establish what they are, or whether satisfaction with Business Link services is a function of performance alone. Consequently the DTI approach cannot assess effectively the inputs to the standards being used by customers so that customer relations cannot be fully understood; (4) they do not assess the impact of the support on the same sample group as that used for assessing advice satisfaction. Critically it is argued here that a satisfaction rating alone does not indicate if the Business Link has helped to improve the competitive position of the firm (thereby achieving one of the Business Link objectives). A proper assessment should be holistic. Finally (5) the DTI assessment methods do not consider in any great depth outcomes of performance measures. Detailed consumer CB and customer feedback to improve the service is neglected and represents a major flaw in the DTI evaluation of this policy. Overall existing surveys and current practice is superficial, provides inaccurate measures of performance, and is an unreliable indicator of the effectiveness of this

government policy.

A wider approach to CS/D & CB can provide comment on (1) whether the expectations, ideal, normative or experienced-based disconfirmation paradigm is being used by Business Link customers, or whether the S/D judgement in this context is more a function of performance alone; (2) if the comparison standards are common across the sample, and how (if evident) the standards are being influenced by marketing inputs; (3) what customers expect, anticipate, or hope will happen as a result of using the Business Link. It should be able to comment on whether they are 'realistic', and reflect what are the commonly held beliefs of why business owner-managers use external support organisations; (4) if there is evidence of mediocre or high variation performance through CS ratings. Which aspects of performance need additional training and whether a general business adviser can provide a satisfactory diagnosis of the contextual operating environment which is critical for correct referral onto associated forms of support; (5) whether a qualitative impact study is effective in securing 'meaningful' data. It should comment on the customers' ability to evaluate the impact of Business Link performance in descriptive and quantifiable terms, identifying from clients those aspects most highly valued; (6) whether dissatisfied customers are complaining. It should establish the redress customers want, whether complaints monitoring is an accurate way of securing constructive customer feedback in this context, and hence whether complaints monitoring can be used to evaluate the policy; and finally (7) detail a range of outcome of performance measures linked to the underlying philosophy and remit of the policy. Only when this full evaluation takes place will reliable indicators of performance be available. These methods are now being applied in further empirical research on Business Link by the author.

REFERENCES

- Adams, J. Stacy (1963), "Towards an Understanding of Inequity," *Journal of Abnormal and Social Psychology*, 67, October, 422-436.
- ADEPT (1995), *Guidelines for Interpreting BS EN ISO 9000 for Business Links*, Produced on behalf of Department of Trade & Industry, SFBLD.

- Anderson, Rolph E. (1973), "Customer Dissatisfaction: The Effect of Disconfirmed Expectancy on Perceived Product Performance," *Journal of Marketing Research*, 10, February, 38-44.
- Bateson, John E. G., and Jochen Wirtz (1991), *Modelling Consumer Satisfaction. A Review*, London Business School, Working Paper Series, 99, August.
- Bennett, Robert J. (1995), *Meeting Business Needs in Britain. Engaging the Business Community through New Style Chambers*, Association of the British Chambers of Commerce, London.
- Bennett, Robert J. (1996), *SMEs & Public Policy: Present Dilemmas, Future Priorities & the Case of Business Links*, Prepared for key note address to 19th ISBA National Small Firms and Research Conference: Enterprising Futures, University of Cambridge.
- Bolton, John E. (1971), *Small Firms: Report of the Committee of Inquiry on Small Firms*, HMSO, London.
- BPRI (1997), *DTI - Business Link Enquiry Services, 1995-6 Mystery Enquirer Programme*.
- Browne, John (1995), *Business Link Best Practice*, Report prepared for the Department of Trade and Industry, London.
- Cardozo, Richard N. (1964), "Customer Satisfaction: Laboratory Study and Marketing Action," *Journal of Marketing Research*, 2, August, 244-249.
- Carma (1995), *BL Monitoring Report*, Report prepared for the Department of Trade & Industry, London.
- Chrisman, James J. and Frances Katrisha (1995), "The Small Business Development Centre Programme in the USA: A Statistical Analysis of its Impact on Economic Development," *Entrepreneurship & Regional Development*, 7, 143-155.
- Churchill, Gilbert A. and Carol Surprenant (1982), "An Investigation into the Determinants of Consumer Satisfaction," *Journal of Marketing Research*, 19, November, 491-504.
- DTI (1992), *Foreword to One Stop Shop Prospectus*, DTI, 1992, 1.
- DTI (1996b), *BL Pilot Evaluation*, Department of Trade & Industry, London.
- DTI (1996a), *Accreditation Booklet*, Department of Trade and Industry.
- DTI (1997a), *Enhanced Business Links - A Vision for the 21st Century*, Department of Trade and Industry, October URN 97/891.
- DTI (1997b), *Management Information*, Schedule Na - Business Enterprise Support Through Business Links, RBM3 (Management Information & Systems Unit), 14.
- DTI (1997c), *Measuring the Impact of Business Link*, Interdepartmental Working Group, Department of Trade and Industry, London.
- Erevelles, Sunil and Clark Leavitt (1992), "A Comparison of Current Models of Consumer Satisfaction/Dissatisfaction," *Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behaviour*, 5, 110-114.
- Ernst & Young (1996), *Evaluation of Business Links*, A report prepared on behalf of the Department of Trade and Industry, January.
- Greysier, S. A. (1976), "Introduction to Conference Proceedings," *Conceptualisation and Measurement of Consumer Satisfaction/Dissatisfaction*, Proceedings of the Conference Conducted by the Marketing Science Institute with support from the National Science Foundation, H. Keith Hunt, (ed.), Marketing Science Institute, Cambridge, Massachusetts.
- HC 302-I (1996), *Business Links*, House of Commons Trade and Industry Select Committee Report, together with the Proceedings of this Committee, London, HMSO, HC 302-I, Session 1995-1996.
- Heseltine, Michael (1992), *Hansard*, 3 December, 1992, 415-424.
- Holbrook, Morris B. (1984), "Situation-Specific Ideal Points and Usage of Multiple Dissimilar Brands," *Research in Marketing*, 7, J. N. Sheth, (ed.), Greenwich, CT: JAI Press Inc., 93-112.
- Hunt, H. Keith (1976), *Conceptualisation and Measurement of Consumer Satisfaction and Dissatisfaction*, Proceedings of the Conference, Conducted by the Marketing Science Institute with support from the National Science Foundation, H. Keith Hunt, (ed.), Cambridge, Massachusetts.
- ISG 13/96 Business Link Implementation Strategy Group (1996), Department of Trade and Industry, BL Directorate, *Meeting papers*.
- ISG 19/96 Business Link Implementation Strategy Group (1996), Department of Trade and Industry, BL Directorate, *Meeting papers*.
- Jacoby, Jacob, and J. James Jaccard (1981), "The Sources, Meaning and Validity of Consumer Complaints: A Psychological Analysis," *Journal of Retailing*, 57, 3, Fall.
- Jayanti, Rama and Anita Jackson (1991), "Service Satisfaction: An Exploratory Investigation of Three Models," *Advances in Consumer Research*, 18, Marvin E. Goldberg, Gerald Gorn and Richard W Pollay, (eds.), New Orleans: Association for Consumer Research, 603-609.
- KPMG (1994), *DTI Evaluation of Business Links*, Department of Trade & Industry.
- La Tour, Stephen A. and Nancy C. Peat (1979), "Conceptual and Methodological Issues in Consumer Satisfaction Research," *Advances in Consumer Research*, 6, W. Wilkie, (ed.), Ann Arbor, MI: Association for Consumer Research, 431-437.
- Mall Research Services (1994), *Business Link Pilot Study - Quality of Service Measurement*, Retail Marketing Consultants, October.
- Mall Research Services (1998), *Business Link Training and Development Survey*, January.
- Miller, John A. (1976), "Studying Satisfaction, Modifying

- Models, Eliciting Expectations, Posing Problems, and making Meaningful Measurements," *Conceptualisation and Measurement of Consumer Satisfaction and Dissatisfaction*, H. K. Hunt, (ed.), Cambridge, MA: Marketing Science Institute, 72-91.
- MORI (1996), *Business Links: Baseline Tracking Study*, Business Advice Among Small and Medium Sized Enterprises and the Business Link Campaign - Wave 4.
- Naumann, Earl and Kathleen Giel (1995), *Customer Satisfaction Measurement and Management*, International Thomson Publishing, USA.
- 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. (1981), "Measurement and Evaluation of Satisfaction Processes in Retail Settings," *Journal of Retailing*, 57, 3, Fall.
- Olshavsky, Richard W. and John A. Miller (1972), "Consumer Expectation, Product Performance and Perceived Product Quality," *Journal of Marketing Research*, 9, February, 179-189.
- Olson, Jerry C. and Philip Dover (1979), "Disconfirmation of Consumer Expectations Through Product Trial," *Journal of Applied Psychology*, 64, April, 179-189.
- Patterson, Paul G., Lester W. Johnson and Richard A. Spreng (1997), "Modelling the Determinants of Customer Satisfaction for Business-to-Business Professional Services," *Journal of the Academy of Marketing Science*, 25, 1, 4-17.
- Pfaff, Martin (1972), "An Index of Consumer Satisfaction," *Proceedings of the Third Annual Conference*, Association for Consumer Research November, 713-737.
- Spreng, Richard A., Scott B. MacKenzie and Richard W. Olshavsky (1996), "A Re-examination of the Determinants of Consumer Satisfaction," *Journal of Marketing*, 60, July, 15-32.
- Titchener, Alan (1996), *Personal Quote*.
- Tucci, Louis A. and James Talaga (1997), "Service Guarantees and Consumers' Evaluation of Services," *Journal of Services Marketing*, 11, 1, 10-17.
- Woodruff, Robert B. and Sarah F. Gardial (1996), *Know Your Customer. New Approaches to Understanding Customer Value and Satisfaction*, Blackwell, Oxford.
- Woodruff, Robert B., Ernest R. Cadotte and Roger L. Jenkins (1983), "Modelling Consumer Satisfaction Processes using Experience-Based Norms," *Journal of Marketing Research*, 20, August, 296-304.

Send correspondence regarding this article to:

Susan J. Priest
University of Cambridge
Department of Geography
Downing Place
Cambridge, CB2 3EN ENGLAND

CUSTOMER SATISFACTION FROM A SUPPLY CHAIN PERSPECTIVE: AN EVOLUTIONARY PROCESS IN ENHANCING CHANNEL RELATIONSHIPS

**Stanley E. Fawcett, Brigham Young University
Michael J. Swenson, Brigham Young University**

ABSTRACT

Meeting customers' real needs is the key to competitive success. Unfortunately, concepts of customer service and satisfaction are frequently misunderstood and often poorly defined--even at leading companies. The result is that customers continue to complain about the poor service that they receive. In fact, articles in the business press indicate that despite extensive efforts to provide ever higher levels of service, customers remain dissatisfied and demonstrate little. Creating true customer satisfaction and achieving sustained loyalty remains a persistent challenge. To better meet this challenge, today's firms need to reconsider their customer fulfillment strategies so that they can keep customers from either defecting to competitors or going out of business. The objective of any new customer-service paradigm must be to clearly define and describe the basic characteristics and likely outcomes of various service activities. As this is done, everyone within an organization comes to better understand exactly what needs to be done to profitably provide customers with a unique and valued set of satisfactions. When everyone within the firm acts on this understanding, the firm is better able to provide truly superior product/service offerings that will help key customers enhance their own competitiveness. This ability to help customers achieve greater competitive success by delivering a value-added capability is the essence of profitable customer takeaway.

INTRODUCTION

Managers at leading firms around the world and across many industries now recognize that the ability to deliver profitable customer takeaway is critical to long-term survival in today's intensely competitive and fast-moving global marketplace. The need to create profitable customer takeaway has been greatly magnified by the fact that channel power has shifted down the supply chain toward

the final consumer in recent years (Blackwell 1997). Shifting channel power has created what has been called the "high-service sponge"--customers that use their market leverage to constantly demand higher levels of service. High-service sponges have a seemingly inexhaustible capacity to "soak up" more of their suppliers' resources to fuel their own quest for market dominance. The emergence of these service-hungry customers who possess tremendous channel power places tremendous pressure on firms throughout the supply chain to develop the capabilities needed to deliver real and valued takeaway.

Firms that fail to deliver valued takeaway set themselves up to be role shifted out of the supply chain--replaced by firms that possess greater and more targeted value-added capabilities. By contrast, firms that consistently deliver high levels of takeaway assure themselves of a secure position in tomorrow's highly integrated supply chains. Unfortunately, many managers do not fully understand the basic nature of profitable customer takeaway. The objective of this article is, therefore, to define profitable customer takeaway, highlighting the underlying themes that make the successful implementation of takeaway-driven initiatives possible. To do this, fundamental distinctions between takeaway and traditional notions of customer service and satisfaction are identified and discussed. Indeed, clarifying the issues--such as organization responsibility, inherent resource requirements, limitations, and influence on customer behavior--that differentiate basic customer fulfillment philosophies is a primary goal of this article.

PROFITABLE CUSTOMER TAKEAWAY DEFINED

Managers have long viewed customer satisfaction as the target result of their firms' customer service activities (Anderson, Fornell and Lehmann 1994; Fornell, Johnson and Anderson

1996). Within this traditional orientation, customer satisfaction is determined by the customer's expectations regarding, and experience with, the product/service package delivered by the firm (Oliver 1980; Spreng, Mackenzie and Olshavsky 1996; Yi 1990; Zeithaml, Berry and Parasuraman 1993). When the customer's expectations are met, confirmation and satisfaction result. When expectations are exceeded, customer delight emerges. However, when expectations are not fulfilled, disconfirmation and dissatisfaction result. From this traditional perspective, the key to achieving high levels of satisfaction is to understand customers' needs so that the firm can develop and deliver distinctive products and services to meet those needs (Oliver 1997). When the firm performs well, its customers take a degree of satisfaction away from the relationship. Unfortunately, in today's global marketplace, which is characterized by intense competition and ever rising customer expectations, delivering this level of takeaway--a degree of satisfaction--is no longer adequate to assure long-term competitiveness.

The Capability Dimension

Empirical evidence demonstrates that even highly satisfied customers exhibit little loyalty, defecting to the competition when circumstance and convenience permit (Fierman 1995, Jones and Sasser 1995, Stewart 1995, 1997). Customer takeaway recognizes this reality and the challenge that it presents--today's successful firms must go beyond satisfying customers to actually enhancing customers' competitive performance. That is, customers take away from the relationship an enhanced capability that enables them to achieve greater market success. Whereas customer-satisfaction strategies seek to make the customer "happy," customer-takeaway initiatives strive to help the customer become a better, more capable competitor in its own markets. By literally helping customers succeed, the firm achieves a certain amount of indispensability since it delivers a capability that is inextricably intertwined with the customer's own competitive success.

With its focus on improving the customer's competitiveness, customer takeaway is the next paradigm of customer fulfillment and possesses the

following unique characteristics:

- Takeaway is truly knowledge based, requiring an intimate understanding of the overall supply chain's competitive imperatives.
- Takeaway relies on the strength of upstream suppliers while working to enhance the competitiveness of the downstream customers; i.e., takeaway is capability based.
- Takeaway not only enhances the competitiveness of each member of the supply chain but makes the ultimate consumer better off.
- Takeaway focuses on system-wide capabilities that make the overall supply chain more competitive. (In a global market where supply chains compete against supply chains, this benefit of takeaway is a vital motivator.)

The Profitability Dimension

The second aspect of "profitable" customer takeaway that merits pointed attention is the notion of verifiable profit. That is, in their effort to provide exceptional service to key customers and achieve high levels of customer delight, many firms have established policies and practices that result in money losing relationships (Bowersox et al 1995). The fact that key accounts are often unprofitable has only been discovered as more rigorous and accurate costing systems have been implemented. Unfortunately, as many firms have begun to reduce their customer base, they have eliminated smaller accounts that were perceived as less important but that were actually profitable.

Achieving profitable customer takeaway requires not only an understanding of the costs required to sustain the relationship but also operational excellence to minimize the costs associated with delivering the value and capabilities being demanded by high-service sponges (Tyndall and Kamauff 1998). To deliver takeaway profitably, the firm must understand and reduce the total landed cost of its product/service offerings. Making the firm's production and delivery systems efficient enough to guarantee profitability is difficult in almost all industries

today and requires that the relevant value-added processes be simplified and made more transparent. Further, because value-added processes span a variety of functional areas within the firm as well as organizational boundaries throughout the supply chain, better information sharing coupled with conscientious efforts to break down the functional and interorganizational barriers are needed to better define appropriate customer takeaway while assuring profitability.

To summarize, profitable customer takeaway is not simply a set of satisfactions. Rather, it is an enhanced competitive capability that leads to long-term market success—for both downstream customers and the entire supply chain. Of course, to be profitable, a high level of operational excellence that is facilitated by other members throughout the supply chain must exist.

MOVING TOWARD PROFITABLE CUSTOMER TAKEAWAY

While meeting customers' real needs has been identified as the key to long-term success (Ohmae 1988), comparatively few companies excel at helping their customers achieve higher levels of competitive performance. Part of this challenge stems from the fact that the related concepts of service, satisfaction, delight, and takeaway continue to be misunderstood (Stock and Lambert 1992). As a result, even leading companies suffer from poorly defined attempts to achieve profitable customer takeaway. This persistent challenge of defining and then achieving profitable customer takeaway suggests the need to revisit the notions of customer service and satisfaction (Bowersox and Closs 1992). Indeed, providing high levels of customer service, especially as service has long been defined and operationalized, is not enough to secure competitive advantage. Customer satisfaction also falls short in a global arena where new products and services are constantly introduced, obsoleting "old" technologies and processes with little or no notice. Even customer delight, the marketer's objective, cannot assure repeat business--much less profitable repeat business--in the current dynamic and intensely competitive environment (Jones and Sasser 1995). Customer success, the ability to improve the customer's competitive advantage by providing

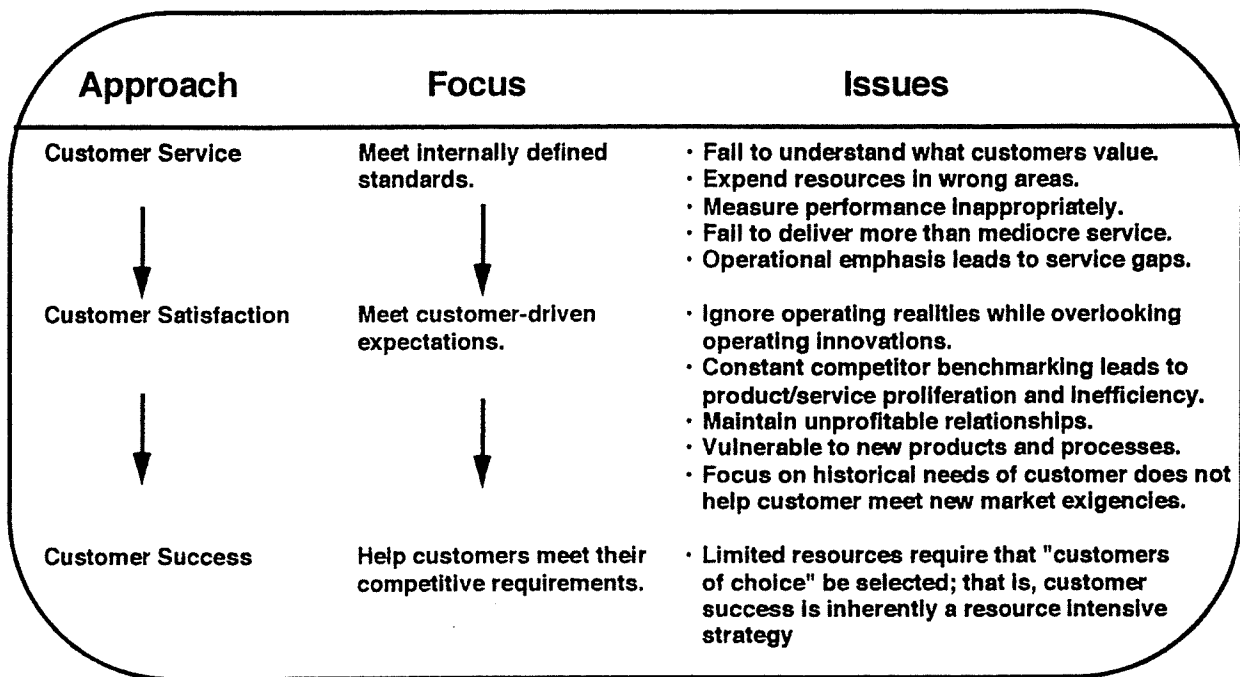
real and profitable customer takeaway, appears to be the firm's ultimate objective. The following paragraphs discuss each of these notions as they relate to buyer-supplier relationships within the supply chain.

Customer Service

Customer service initiatives have historically focused on meeting internally defined standards as they relate to what the firm views as important activities or processes (see Table 1). Measures of customer service therefore take the form of percent defective products, percent of jobs finished on time, fill rate from distribution centers etcetera. By performing well in these areas, the firm hopes to serve the customer adequately. Unfortunately, a common feature of these measures is that they are internal to the firm and are measured by an internal auditing system. Further, because firms use a multitude of measures that account for wide-ranging issues related to cost, time, and performance; managers and workers alike often begin to feel that they are achieving high levels of desirable service even when they are not. The inward focus of the definition and the accompanying measures often lead the firm to overlook the imperative of understanding the customer and meeting the customer's real needs. Far too frequently, firms substitute massive efforts and resources for the well-tailored customer service programs required to deliver profitable customer takeaway. Without systematic customer feedback, it is too easy to emphasize the wrong service activities or build the wrong products and thereby dissipate tremendous resources on becoming excellent at something that is not valued by the customer (Stock and Lambert 1992).

Appropriate customer feedback can help answer the following important questions. How does an important customer define quality? How does our customer measure quality? Is our internal measure of quality consistent with the customer's measure of quality? Does our current level of 99 percent internal quality meet customer requirements? Would an improvement in our quality levels really be valued by our customer? Would an improvement in our quality levels really increase the competitiveness of our customer? Without this type of external focus, a firm's

Figure 1
Customer Service, Satisfaction, and Success Fundamentals



management team is placed in a position of hoping that it has selected the correct activities and measures to achieve the exceptional service desired by the customer. When the wrong activities or measures are used, both mediocrity and frustration result. For example, a division of one manufacturing company set quality performance standards at a level lower than the customer's (a sister division) expectations. Shipment after shipment that passed the internal standards was returned as unacceptable. Aligning quality standards at the higher level would have required additional training and investment but would have lowered long-term costs and eliminated substantial frustration and intra-firm rivalry. Discrepancies of this kind occur on a frequent basis when the firm operates with a traditional inward-looking customer service philosophy that emphasizes efficient, and even excellent, operations over appropriate understanding and aligned operations. Such discrepancies result in service gaps, which are the equivalent of an open invitation for competitors to enter the market and "steal" valuable customers (Parasuraman, Zeithaml and

Berry 1985).

Customer Satisfaction

Customer satisfaction initiatives recognize the threat that service gaps represent. The focus engendered by satisfaction programs is therefore on obtaining direct input from important customers regarding their service expectations. The goal is to eliminate service gaps by meeting customer-defined expectations better than the competition (Zeithaml, Parasuraman and Berry 1990). Thus, to the same extent that customer service looks inward, customer satisfaction focuses outward. Likewise, just as customer service measures are internal to the firm, satisfaction measures are externally oriented, requiring customer feedback. The impact of customer input can vary substantially. Customer input might simply help the firm modify existing measures so that they are in better alignment with customer expectations, or it might lead the firm to reallocate resources and reevaluate priorities, or it might motivate the firm to adopt entirely new policies or practices. The

critical issue is that customers are more centrally involved in defining and evaluating the performance relationship between the two firms. In other words, achieving high marks in customer satisfaction requires that the firm learns to benchmark against customer requirements

Several potential challenges emerge when the emphasis is on maintaining satisfied customers. Perhaps the most common challenge associated with a strong desire to satisfy customers is that some managers make promises that cannot be fulfilled and that ultimately lead to both alienated customers and considerable operating conflicts and confusion within the firm. Alternatively, in a desire to outdo competitors, service and product offerings can easily proliferate with a resultant loss of efficiency and perhaps relevancy. Relatively few customers have the foresight to turn down new products or services that might meet a niche need without creating real value or expanding real market opportunities. A close corollary occurs when the firm decides that it must meet the customer's requirements at any cost. This scenario might lead to high levels of customer takeaway without contributing to the firm's profitability and long-term survival. Finally, even highly satisfied customers can go out of business, leaving a firm with a diminished customer base. A satisfied customer is not always a successful customer, and successful customers are needed for sustained operations. This limitation to customer satisfaction initiatives arises when important customers become complacent or inward looking and lose sight of the value they really need to create for long-term success. Most of the limitations of a customer satisfaction approach diminish efficiency and the profitability of the buyer/seller relationship and the overall supply chain. Too heavy a focus on the customer and what has worked in the past can leave the relationship vulnerable to the exigencies of a dynamic marketplace and the predatory practices of global competitors.

Customer Success

Customer success initiatives are founded on the understanding that sustainable customer competitiveness is more important than maintaining a set of customers who are currently satisfied.

Such initiatives also implicitly recognize that the buyer/supplier relationship must yield competitive advantage, or profit, to both firms. In fact, in a world where supply chains compete against each other for world market share, customer success endeavors focus on building a supply chain team that possesses world-class capabilities. Bringing these ideas together highlights the two fundamental distinctions between achieving satisfaction and success. First, helping customers achieve success requires a knowledge of the entire supply chain. That is, whereas customer satisfaction strategies require information to bridge the gap between the firm and its customers, customer success requires that a firm understand what its customers' customers desire in a product/service package. This knowledge of downstream requirements is needed to provide a "better" product/service mix--a mix that leads to enhanced customer competitiveness. It also helps to focus the entire supply chain on creating the greatest possible competitiveness for the entire supply chain team. One CEO explained the essence of creating customer success as follows, "We turn our customers into winners. Their success is cash in our bank. Our customer is our most important partner in cooperation--his customer benefits from this as well." (Ginsburg and Miller 1992).

Turning customers into winners can be a difficult process both in gaining the requisite knowledge and expertise as well as in translating that expertise into competitive advantage for the customer. The firm, and its supply chain partners, must invest in specialized skills to obtain vital information about downstream requirements that can then be used to tailor the firm's product/service packages to deliver exceptional value. At the same time, the firm takes on the role of consultant to its customers, educating them in areas where they lack needed skills or knowledge. When a good relationship exists between the firm and its customers, this educational role goes relatively smoothly and is genuinely appreciated. One manager pointed out that a particular supplier was indispensable because the supplier knew more about the industry than his firm, and more importantly, actively shared that knowledge to help his firm be more successful. However, when relationships are not well established, informing customers that they perhaps

"are not always right" can be uncomfortable. Even so, leading firms realize that it is better to turn down a customer request for a product or service when the product or service in question is inappropriate for the customer's real needs. The key is to help the customer identify and/or develop a product/service package that will more successfully fulfill market requirements and enhance the customer's competitiveness.

Second, profitably enhancing customer performance requires an integrative approach that balances internal operations and external satisfactions. Meeting the real needs of the high-service sponge yields the firm very little advantage if the firm cannot do so at a profit. Of course, legitimate occasions exist when a firm might incur a short-term loss in order to secure a market relationship that will prove profitable over the duration of the relationship. The challenge is twofold: first, to identify attractive opportunities where the firm's capabilities match the market needs of key customers and second, to coordinate the firm's--and its supplier's--value-added efforts so that the firm can efficiently deliver valuable takeaway to the selected customers. Even as management begins to evaluate market opportunities, a need exists to integrate marketing and operations activities to better define potential takeaway and determine whether the needs of specific customers can really be met efficiently and effectively. This evaluation process necessitates that managers realize that not all customers are equal and that individual customers have different needs that require different resource commitments. Managers must select "customers of choice" and then help them succeed by delivering value-added services that are unique to them.

Having identified key customers and what constitutes real takeaway for them, the task is to achieve the communication and coordination necessary to use the diverse value-added activities of the entire upstream supply chain to profitably create and deliver the required set of takeaways. To do this, takeaways must be clearly defined and communicated and each part of the organization as well as each member of the supply-chain team must understand 1) its role in the value-added process, 2) the value that is added by other areas within the firm and across the supply chain, and 3) how decisions made in one area impact other areas

as well as the overall value-added process. This level of understanding provides the foundation for cross-functional and inter-organizational integration and process management. Efficient and effective takeaway results when synergistic integration is sought after and achieved.

CONCLUSIONS AND IMPLICATIONS

In today's dynamic global marketplace, the firm must use a set of constrained resources—its own combined with its supply chain allies'--to help key customers go beyond satisfaction to competitive success, and to do so better than a constantly evolving and improving set of global competitors (Drucker 1994). The notion of profitable customer takeaway, which can appropriately be defined as providing product/service packages that help customers succeed while yielding profit to the firm, can help a firm build a world-class supply-chain team and a sustainable competitive advantage. The critical need in getting to this point is to carefully evaluate the firm's approach to customer relations to determine whether or not it is conducive to the creation and delivery of profitable customer takeaway. The customer service, satisfaction, and success paradigm described above provides the firm a framework to explicitly evaluate the level of profitable customer takeaway that is being provided. By evaluating and answering the following questions, managers can help assure the success of both their firm and the overall supply chain.

- Does the firm still operate from a service orientation that is driven by a set of internal measures of activity performance? If so, are the takeaways provided the ones that customers truly want, and does the firm provide them with world-class efficiency?
- If the firm has adopted a true customer satisfaction orientation, are the systems and measures in place to really know what customers want and need? Or is the firm operating on past assumptions? Further, has the emphasis on customer satisfaction adversely impacted operations in a way that hurts profitability?

- Does the firm really understand the supply chain and the competitive needs of the different members of the supply chain? What imperatives are changing the competitive dynamics of the supply chain?
- Does the firm know how profitable different customer relationships are? Where are the capabilities created that lead to enhanced customer, and supply chain, performance?
- Could specific value-added roles be better defined and more appropriate processes developed? Could communication and coordination be improved to deliver more profitable customer takeaway? If so, how?
- Is the firm providing value to other supply chain members that will enhance the overall competitiveness of the supply chain? Can the firm's product/service package be provided by another member of the supply chain in a way that threatens the firm's active participation through role shifting?

REFERENCES

- Anderson, E. W., C. Fornell and D. R. Lehmann (1994), "Customer Satisfaction, Market Share, and Profitability: Findings from Sweden," *Journal of Marketing*, 58, January, 53-66.
- Blackwell, R. D. (1997), *From Mind to Market: Reinventing the Retail Supply Chain*, New York: HarperBusiness.
- Bowersox, D. J., R. Calantone, S. Clinton, D. Closs, M. Cooper, C. Droge, S. Fawcett, R. Frankel, D. Frayer, E. Morash, L. Rinehart and J. Schmitz (1995), *World Class Logistics: The Challenge of Managing Continuous Change*, Oak Brook, IL: Council of Logistics Management.
- Bowersox, D. J. and D. J. Closs (1992), "Adding Value by Distribution," *Logistics Technology International*, 218-221.
- Drucker, P. F. (1994), "The Theory of the Business," *Harvard Business Review*, 72, 5, 95-104.
- Fierman, J. (1995), "Americans Can't Get No Satisfaction," *Fortune*, December 11, 186-194.
- Fornell, C., Michael D. Johnson and Eugene W. Anderson (1996), "The American Customer Satisfaction Index: Nature, Purpose, and Findings," *Journal of Marketing*, 60, October, 7-18.
- Ginsburg, I. and N. Miller (1992), "Value-Driven Management," *Business Horizons*, 77, 3, 23-27.
- Jones, T. O. and Jr. W. E. Sasser (1995), "Why Satisfied Customers Defect," *Harvard Business Review*, 73, 6, 88-99.
- Ohmae, Kenichi (1988), "Getting Back to Strategy," *Harvard Business Review*, 66, 6, 149-156.
- Oliver, R. L. (1980), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, 17, November, 460-69.
- Oliver, R. L. (1997), *Satisfaction: A Behavioral Perspective on the Consumer*, New York: McGraw-Hill.
- Parasuraman, A., V. A. Zeithaml and L. L. Berry (1985), "A Conceptual Model of Service Quality and Its Implications for Future Research," *Journal of Marketing*, 49, 4, 41-50.
- Spreng, R. A., S. B. Mackenzie and R. W. Olshavsky (1996), "A Reexamination of the Determinants of Consumer Satisfaction," *Journal of Marketing*, 60, July, 15-32.
- Stewart, T. (1995), "After all you've done for you customers, why are they still NOT HAPPY?" *Fortune* December 11, 178-18.
- Stewart, T. A. (1997), "A Satisfied Customer Isn't Enough," *Fortune*, July 21, 112-113.
- Stock, J. and D. Lambert (1992), "Becoming a "World Class" Company With Logistics Service Quality," *International Journal of Logistics Management*, 3, 1, 73-80.
- Tyndall, G. R. and G. W. Kamauff (1998), *Global Supply Chain Management*, New York: John Wiley & Sons, Inc.
- Yi, Y. (1990), "A Critical Review of Consumer Satisfaction," *Review of Marketing*, V. A. Zeithaml, (ed.), Chicago: American Marketing Association, 68-123.
- Zeithaml, V., A. Parasuraman and L. L. Berry (1990), *Delivering Service Quality: Balancing Customer Perceptions and Expectations*, New York: The Free Press.
- Zeithaml, V. A., L. L. Berry and A. Parasuraman (1993), "The Nature and Determinants of Customer Expectations of Service," *Journal of the Academy of Marketing Science*, 21, Winter, 1-12.

Send correspondence regarding this article to:

Stanley E. Fawcett
 Marriott School of Management
 Brigham Young University
 Provo, UT 84602 USA

AN INTEGRATIVE, UTILITY-BASED APPROACH TO PROFITABLE CUSTOMER TAKEAWAY IN THE SUPPLY CHAIN

Stanley E. Fawcett, Brigham Young University
Michael J. Swenson, Brigham Young University

ABSTRACT

Integrating functional areas of the firm to address supply chain decisions and activities is critical for firms seeking to deliver profitable customer takeaway. In this paper, the domain of profitable customer takeaway is clarified and expanded to reflect customer capabilities. Value-added activities and processes that create utilities for customers in the supply chain are presented. The utilities classification framework captures the complexities of customer takeaway in the supply chain and illustrates resulting strategies for functional integration.

INTRODUCTION

Because of the growing intensity of competition in virtually every sphere of business activity, managers are increasingly aware of the need to better understand and integrate functional activities in the supply chain. Anderson and Narus (1999) observe that "firms migrate to a supply management orientation when managers realize that the fate of their organization is inextricably linked to other companies in the value network." Blackwell (1997, p. 126) notes that "partners in winning supply chains will be expected to add value and efficiency to the chain or jeopardize their position in that chain." Patterson, Johnson, and Spreng (1997) recognize that customer satisfaction/dissatisfaction "is the crucial link in establishing longer-term client relationships and thus the strategic well being of the organization."

Yet firms struggle in their quest to provide profitable customer takeaway in the supply chain for at least two reasons. First, most firms continue to possess a myopic view of what customer takeaway entails. They do not yet realize that they have a vested interest in making their customers successful. Second, many firms fail to properly define value creation in terms of real customer capabilities. They then exacerbate the problem by creating functional and inter-organizational barriers that inhibit the creation and delivery of valued capabilities (Fawcett and

Fawcett 1995). The bottom line is that many firms not only fail to do the right things to help their customers succeed but they also dissipate profitability through awkward and inefficient capabilities.

To position the firm to deliver profitable customer takeaway, managers must return to the basics; that is, to the essence of what a customer needs from its suppliers as well as the value it must deliver to its own customers in the supply chain. The four fundamental economic utilities—form, time, place, and possession—provide useful insight that can help managers get back to the basics of facilitating customer success. Indeed, these utilities highlight the rationale behind 1) the acquisition of products or services and 2) the development of buyer-supplier relationships. Quite simply, companies enter into a supply chain relationship because of the utilities or set of satisfactions and capabilities they expect to receive. Form, time, place, and possession utilities embody the value and capabilities that companies seek to acquire and use in their quest for success. By establishing the capabilities required to 1) build a better product, 2) deliver it on time to the right place, and 3) enable the customer's own creative ability, today's firm assures itself of enhanced competitive success.

Unfortunately, consistently creating utility for customers is a daunting task that requires managers to understand and to manage the activities and processes involved in the product/service/customer interaction. It is through the customer's interface with a product/service package that the customer gains some form of enhanced competitive capability. Whether this competitive capability comes from the use of a technologically advanced product (Intel inside) or the acquisition of a specialized service (3rd-party logistical management), it can be neither developed nor shared with supply chain partners without the meticulous management of a diverse set of activities and processes. Knowing how these activities and processes interact and where they reside within the organizational structure of the firm is vital to managing them for the achievement

Table 1
Takeaway, Utilities, and Functional Interactions

Takeaway	Activities/Processes	Functional Areas
Possession Utility	<ul style="list-style-type: none"> • Recognize the firm's distinctive capabilities • Identify and evaluate customers • Understand supply chain imperatives • Define customer success factors • Communicate success factors throughout the firm • Select customer of choice • Build relationships based on profitable customer takeaway 	<ul style="list-style-type: none"> • Operations leads cross-functional effort • Marketing leads cross-functional effort • Marketing, Operations, Logistics, Purchasing • Marketing takes lead • Top management • Marketing, Operations, Logistics • Marketing takes lead
Form Utility	<ul style="list-style-type: none"> • Balance market pull and technology push pressures in developing distinctive capabilities • Focus on selected capabilities and respective processes • Convert raw materials and component inputs into finished product/service package 	<ul style="list-style-type: none"> • Engineering provides technical expertise while Marketing provides market awareness • Operations takes lead; Purchasing supports • Operations creates the form
Time Utility	<ul style="list-style-type: none"> • Make products and services available to customers when they want them 	<ul style="list-style-type: none"> • Purchasing provides materials on time; Operations meets production due-dates; Logistics stages and moves products; and Marketing interacts with customer
Place Utility	<ul style="list-style-type: none"> • Make products and services available where customers want them 	<ul style="list-style-type: none"> • Logistics and Marketing

of maximum takeaway.

The primary objective of our article is to take an initial step in bridging the gaps between customer takeaway, functional integration, and profitability in the supply chain. In the next sections, we present a variety of value-added activities/processes that must be performed to create utility. Among the various types of utility that economists classify, we focus on four: possession, form, place and time. Furthermore, we discuss the specific roles different functional areas must play to make the activities/processes both effective and efficient. This presentation is summarized in Table 1.

Possession Utility

Possession utility has traditionally emphasized the dissemination of information to establish both an awareness of a product and a desire to purchase or "possess" the product. From this perspective,

possession utility has been managed quite differently from the other utilities since it does not result from a change in product form or availability. Rather, possession utility emerges from the customer's perceptions regarding pride of ownership and ego involvement. Given this traditional view of possession utility, marketing has taken primary responsibility for its creation. However, in a dynamic marketplace where channel power has shifted toward the consumer and where higher levels of takeaway must be delivered to selected customers, a more comprehensive view of possession utility is warranted.

Today, possession utility must be founded on a clear conceptualization of profitable customer takeaway. Properly defining profitable customer takeaway requires that the firm understand its own unique abilities as well as the real needs of its customer and of the entire supply chain. In effect, creating possession utility requires that the firm develop a viable capability that enables the

customer to do something significant and unique. The firm must offer customers access to a distinctive capability—the best technology available worldwide, the shortest-cycle order fulfillment times, the lowest-cost logistics services, or the highest-quality components or subassemblies. Possession utility can even be the sure knowledge that the firm is a member of the best supply-chain team in the marketplace and therefore possesses the marketing mix (product, price, promotion, place) best suited to meet the ultimate customers' needs. In a world of outsourcing, possessing and sharing a world-class capability with other members of the supply chain assures greater success for the entire supply chain and helps position a firm as both a supplier and customer of choice. With this understanding, the firm is in a position to build better relationships with potential customers and effectively communicate to them the information needed to create both an awareness of and a desire to possess the firm's product/service package. In fact, at times, the communication is directly to the customer's customer in an effort to create awareness and demand throughout the downstream supply chain (consider Intel's branding and promotion of its 386, 486, and Pentium processors). Accordingly, each of the following activities play an important role in achieving a high level of possession utility.

- A concise understanding of the firm's distinctive capabilities is obtained.
- Customers are identified and evaluated.
- A real understanding of supply chain imperatives is gained.
- Customer success factors are explicitly defined and communicated throughout the firm.
- Customers of choice are selected based on an alignment between customer success factors and the firm's distinctive capabilities.
- Strong relationships are built with important customers--relationships are based on profitable customer takeaway.

A firm's efforts to create possession utility begin with an internal focus--an effort to discover what the firm really does well and what the firm's product really is and does [Ohmae, 1988]. Knowing well the firm's distinctive capabilities

enables managers to better define the potential target market. Delimiting the firm's core competencies is a truly cross-functional endeavor, one in which engineering, production, distribution, marketing, and corporate strategy all come together to define and discuss the value that the firm creates and delivers (Prahalad and Hamel 1990). Of course, each group brings a different set of perceptions to the discussion, but each perspective is valuable for at least two reasons. First, as the perspectives are shared openly, each area of the firm gains a better appreciation for what the other areas of the firm must do for the overall organization to succeed. Second, at some point during the discussion, an intersection of perspectives generally reveals the firm's true critical capabilities, highlighting the complexity and cross-functional nature of uniquely distinctive capabilities (Drucker 1994; Stalk et al 1992). From this discovery and analysis effort, an endeavor that often involves the mapping of key processes to enhance their visibility, specific roles are better defined, interactions are understood, and the foundation for greater cooperation is established.

The second important activity requires that managers look outward to the customer. The existing customer base as well as all potential new customers should be carefully identified. Each of the customers should then be carefully evaluated and classified on an "ABC" basis with the "A" customers having the greatest potential importance to the firm. In firms where customer analysis is systematically performed, the identification and evaluation process almost always falls under the domain of marketing. However, the requirements of profitable customer takeaway suggest that while marketing should probably provide leadership throughout this process, other functional areas need to become more involved in the analysis. Marketing is generally in the best position to identify customer needs and determine existing levels of customer satisfaction; but even here, in the area of need definition and satisfaction evaluation, other domains are often able to enhance definitions and measures of satisfaction. In one supply-chain relationship, getting logistics managers actively involved in the evaluation of needs led to the development of a new performance measure and higher levels of

customer success. Where traditional fill rate measures had been used, a better knowledge of customer needs led to the development of a measure that showed the percentage of the time the product was actually on the shelf (after all, regardless of the fill rate from the warehouse, the end consumer cannot buy the product if it is not on the shelf at the retail store level). The measure was then used to enhance communication and modify the relationship between the two firms.

Equally important to understanding the customers needs is the task of determining whether or not those needs can be met profitably. Issues surrounding product/service complexity and customization, production volumes, relationship duration and intensity, and resource dedication all need to be considered from an operating perspective. Sophisticated total costing, activity-based costing, life cycle costing, and contingency analysis should also be performed to better understand the potential for developing a profitable relationship with each major customer. Effective customer analysis thus requires marketing, operating, and costing analyses and participation from multiple functional areas.

Closely related to customer identification and evaluation is the notion of supply chain analysis. To truly provide high levels of customer takeaway for those firms that are viewed as potentially important and profitable, the firm needs to understand what the customers' customers and other downstream supply chain members define as important takeaway. Thus, it becomes important to determine what imperatives drive the success of the supply chain. This form of supply chain analysis is particularly important in today's global operating environment where competition is increasingly "no longer company vs. company but supply chain vs. supply chain." (Henkoff 1994) At IBM, salespeople focus on understanding success factors in particular industries. For example, salespeople who work extensively in the banking industry take classes at the Wharton School of Business in finance and banking so that they understand the needs and circumstances of their customers. The objective is to help salespeople understand customer success factors better than the customer (Yarbrough 1996). For most firms, a cross-functional approach incorporating engineering, operating, and

marketing expertise would be best suited to a comprehensive analysis of supply chain imperatives. In particular, the analysis of the needs of the customer's customers represents an ideal opportunity to utilize the firm's marketing research capabilities.

When customers' perceived needs are combined with a knowledge and understanding of supply chain imperatives, managers are well positioned to clearly and accurately define essential customer success factors for each potential key account. Customer success factors are those distinctive elements that help provide the customer with an enhanced competitive ability. For instance, in the automobile industry, product development lead times have been dramatically reduced such that world-class car makers bring new cars to the market in somewhere between 18 and 30 months (a few rare examples of cars going from concept to market in as little as 15 months exist). In this environment, a parts supplier that can bring new technology and design expertise to a design for manufacturing team can have a substantial impact on its customers' competitive ability. Likewise, for a manufacturer that operates multiple facilities across widely dispersed geographic regions, a logistics service provider that could move raw materials, sub-assemblies, and finished goods seamlessly and efficiently around the world with state-of-the-art information systems and an ability to redirect shipments with little or no advance warning would be an invaluable asset. In general, customer success factors fall into one of the following categories: quality, cost, responsiveness/flexibility, delivery/dependability, innovation, information, and reach. It is important to note that once important customer success factors have been identified, they must be communicated throughout the organization to anyone and everyone that has anything to do with the customer. Not until everyone understands the customer success factors can the resources and commitment needed to deliver the necessary takeaway be mobilized.

When managers understand the firm's distinctive advantages, the needs and attractiveness of potential customers, and the requirements of the supply chain, they can select customers based on the probability of achieving high levels of profitable customer takeaway. That is, where

Figure 1
The Alignment Matrix

High	<p>Efforts to pursue attractive activities for which the firm has no advantage result in diminished focus and dissipated capabilities.</p>	<p>Effective alignment results in profitable customer takeaway.</p>
Customer's Success Factors	<p>Low customer priority and low firm competence means that most firms avoid these activities. Resources expended here are wasted.</p>	<p>Significant effort and resources dedicated to non-valued activities results in low customer takeaway.</p>
Low	Low	High
	Firm's Distinctive Capabilities	

alignment exists between the firm's distinctive capabilities and the customer's important success factors, opportunities to build strong and profitable relationships abound. Figure 1 matches customer success factors and the firm's distinctive capabilities, indicating that the likely outcomes would depend on the degree of alignment. Clearly, the best situation occurs when customers need what the firm does well. In each of the other quadrants, diminished focus, wasted resources, and limited or unprofitable customer takeaway is the likely result.

For most firms, the selection of "customers of choice" where intensive relationships are to be built is a strategic decision. At the very least, senior management establishes guidelines to direct the selection and facilitate the management of these key accounts. At this point, once customers have been selected and the relationship needs evaluated, efforts are directed at developing the appropriate relationships. Considerable time and resources are typically expended to build strong relationships with the most important customers of choice. In these tightly coupled relationships, linkages

between the two firms are established through marketing as well as operations, engineering, and distribution. Of course, not all relationships are viewed as equal in importance and fewer resources will be dedicated to some relationships where limited profitable takeaway is the likely result. The critical issue is that the analysis has been done to determine both whether profitable customer takeaway is probable and what kind of relationship is needed to achieve success. In many instances, a less intensive relationship based on delivering good customer service will be the most appropriate and the most profitable.

Possession utility today involves the matching of specific capabilities to customer success factors and building the necessary supply chain relationships to deliver the identified and promised takeaway. Because the creation of possession utility is knowledge driven and because the needed knowledge--knowledge of capabilities, customer attractiveness, and supply chain imperatives--resides in different areas of the firm, possession utility is inherently cross-functional. Even so, different functional areas will take the lead in

performing the different activities described above that become the foundation of possession utility. For example, engineering, operations, and logistics would likely guide the analysis of core capabilities while marketing would provide leadership in customer analysis. It is even probable that marketing would coordinate all of the activities involved in aligning capabilities to success factors.

Form Utility

Form utility is created when the firm converts or transforms a set of inputs from their existing level of desirability to a higher level of desirability. This conversion or transformation process has traditionally been the domain of the operations function with input from marketing and engineering. However, a greater emphasis on matching distinctive capabilities to customer success factors in order to increase profitable customer takeaway suggests the need for a more integrative approach to creating form utility. For example, in an industry where supply chain competition has led to compressed product life cycles and the need for shorter development lead times (customer success factor), a high level of manufacturing, marketing, engineering, and even purchasing interaction is needed to offer a technologically appropriate and market-acceptable product (distinctive capability). Similarly, to offer real and unique takeaway to customers, a firm has to have some form of distinctive core competence that is valued by the market. Achieving this level of competence generally requires greater focus within the firm and thus greater reliance on other supply chain members. Thus, a more integrative approach to managing upstream suppliers is needed. These two examples illustrate the very close linkage between the firm's core capabilities and form utility. Indeed, most of the activities that must take place to assure maximum form utility have something to do with the selection and development of these capabilities. Three relevant and generalizable activities are discussed below. They are 1) balancing market pull and technology push pressures in developing distinctive capabilities, 2) increasing focus on selected capabilities by outsourcing non-critical activities/products, and 3) achieving an appropriate mix of product and service in the product/service

package provided to selected customers.

First, what constitutes a distinctive, value-added capability is determined by a delicate interplay between what the market perceives as its needs and the development of new product and process technologies. The need for balance comes from the fact that products/service packages that currently meet market needs can be easily obsoleted by new technology. For example, compact audio disk players quickly replaced the phonograph while Wal-Mart stores forced many small-town hardware stores into bankruptcy. Likewise, new technologies without current application or the potential to alter market behavior lack utility. For example, digital audio tape players failed to capture consumers' attention as did McDonald's Arch Deluxe sandwiches. The need for market acceptance and technological rigor requires much greater interaction among engineering, marketing, and operations in the design and development of the firm's products and processes. Engineering brings the technical expertise to the discussion while marketing brings the market awareness. While operations often gets caught between these two, it must translate the ideas of both into a deliverable product/service package. When the appropriate integrative mechanisms are put in place and each area values the other areas' contributive ability, distinctive capabilities that lead to profitable customer takeaway result. The need for cooperation and coordination is a reality whether the selected capability is simultaneous engineering, world-class manufacturing, outstanding branding and promotion, or global distribution.

Second, successful firms generally choose certain capabilities and then dedicate their efforts and resources to their development. That is, limited managerial and financial resources mean that a firm "cannot be all things to all customers" and therefore must focus its scarce resources on what it does best to serve those customers that are most important. This imperative has led many firms to outsource a greater percentage of the products and services needed in their operations (Quinn and Hilmer 1994). The value of purchased inputs is now approximately 60 percent of the typical manufacturing firm's cost of goods sold and frequently up to 30 percent of the costs of a service organization. Equal in importance to the

high percent of operating costs represented by purchased inputs is the fact that these inputs increasingly represent higher valued goods and services that are central to the firm's competitiveness. As a result, the purchasing function has taken on a much more visible and impactful role in the transformation process. Supplier involvement in new product development efforts has also increased the impact of purchasing on the creation of form utility. While the role of purchasing has changed the most in this outsourcing environment, the other areas of the firm must be actively involved in evaluating activities, processes, and product lines to determine which should remain in house and which should be outsourced. The other areas should also play an active role in evaluating the performance of the outside materials and service providers. Finally, depending on the nature of the capabilities that are selected as central to the firm's competitive efforts, one or more functional areas needs to take proactive ownership of their development while drawing on the assistance and expertise of the other functional areas to enhance the capabilities' performance.

Third, helping the customer perform at higher levels often requires the firm to carefully consider the functionality and nature of the entire product/service package being delivered (Ohmae 1988). Perhaps more than ever before, physical products and intangible services have become intertwined to the point that the two must work cohesively together to yield profitable customer takeaway. Take the case of a sophisticated mechanical component to be used in the customer's main product line. Not only does the customer rely on state-of-the-art engineering design and high-quality, low-cost manufacturing but also on superb after-sales service and strong warrantee coverage. A failure in any of the services or in the product itself diminishes the value of the customer takeaway, tarnishing the customer's reputation. The supplying firm's profitability is also hurt as it must cover expensive support services and warrantee costs (the loss of future sales is also a distinct possibility). Another unique example of the inseparability of the product from the service that accompanies it was illustrated by Aisin Seiki Co., a supplier of brake parts, when one of its factories burned down. The factory was

the only source of a valve used in most of Toyota's auto production, and Toyota only kept four hours worth of inventory on hand. Aisin quickly mobilized its management and design teams to get the part back in production in other facilities, including those of competing auto parts suppliers. As a result of this quick action, Toyota's auto assembly plants were back on line within a week and lost production was recouped within a couple of months. Aisin's willingness to go to extreme measures to assist Toyota emerged from the fact that Aisin's own success depends on its ability to provide Toyota with a strategic capability—the design, manufacture, and delivery of an important subassembly. By providing Toyota with an important capability in the form of a combined product/service package, Aisin has become more than a favored Toyota supplier—Aisin has become an extension of Toyota's vaunted manufacturing capability.

Delivering this level of augmented product necessitates the proactive involvement of employees throughout the firm as well as the input of diverse functional areas including engineering, logistics, marketing, operations, sales support, and strategic planning. Proactive employee participation in the design and delivery of the firm's product/service package leads to continual improvements in the firm's products and the processes used to make and deliver them. However, active participation only occurs when the firm invests in adequate training and works to build high levels of motivation and trust. This is particularly true for those employees that come into contact with the customer—their proactive and positive participation is needed to ensure a successful customer/product interaction. Additionally, periodic reviews of the takeaway provided by the firm's product/service package can help assure that key customers' success factors are supported. A cross-functional team with representatives from each functional area that has responsibility over one or more aspects of the product/service package should be given this task. Ultimately, form utility that leads to profitable customer takeaway is the result of both intensive interaction among a diverse set of functional areas and the active involvement of employees throughout the firm.

Place and Time Utilities

Place and time utilities can be jointly discussed because a firm's product/service package only has potential to improve the customer's competitive position if it is where the customer needs it--place utility--when the customer needs it--time utility. These two utilities almost always work together. A significant challenge to providing high levels of time and place utility is that they are frequently overlooked because they are often the product of functional areas that are viewed as support or "satisficing" functions. Logistics, for example, is often charged with moving and storing materials, assemblies, and finished products among the firm's diverse operations as well as between the firm and other supply chain members. Yet, logistics is typically managed as a non-strategic cost center. Unfortunately, companies generally come to appreciate the value of place and time utility--and the functional areas responsible for their creation--only after unfortunate and costly events such as when materials are unavailable and the production line has been shut down or after a product is sold out and customers have opted for a competitor's offering instead. More than one firm has demonstrated an ability to create form utility but failed in the marketplace because it was unable to provide time and place utility. This phenomenon has become even more pervasive in today's global marketplace within which firms have rationalized their manufacturing networks and where firms must deal with a multiplicity of complex and poorly understood distribution channels in order to get the product to consumers worldwide.

The interconnection between time and place utilities can be seen by considering some of the decisions and activities that are basic to their creation. Some of the critical decisions are where, how much, and when to both produce and store as well as by which mode and which carrier to transport. The decision of where to produce a product goes beyond seeking a low-cost, high-quality production location to take proximity to key customers into account and thus affects both place and time utilities. For example, Tijuana has become the world's television manufacturing capital because it combines low-cost, quality-conscious Mexican labor with close proximity to the large U.S. consumer market. Toyota City is

well-known in the automobile industry because of the large number of auto parts suppliers that have co-located next to Toyota's design and assembly facilities. The close location allows for relatively easy just-in-time delivery of components. Thus, where a product is produced impacts how quickly it can be delivered. These production location decisions, including the issue of production scale are generally strategic in nature, but the input of production and marketing managers is generally included. A recent study also indicated that while logistics is generally not a part of this decision, including insight from logistics managers would greatly improve the design and performance of multi-facility networks (Scully and Fawcett 1994).

The related issue of how much to produce involves an initial decision regarding the scale of the operation together with periodic decisions regarding the size and timing of a given production run. Each of these latter decisions impact the time availability of the product and are typically made by production managers. Interestingly, one of the frequent conflicts that arises in manufacturing environments relates to the timing of production--marketing promises product delivery without consulting manufacturing. A frequent result is that manufacturing is forced to expedite orders or miss promised delivery dates. Greater coordination between the two could avoid this problem and enhance both time and place utilities. An extension of this conflict involves logistics. That is, when production misses a promised due date, logistics is called on to expedite delivery, frequently using air freight. The result is that logistics costs escalate dramatically (this is troubling if logistics is evaluated as a cost center). One global manufacturer that planned on using ocean container shipping as the standard mode of transportation found that it had to air freight almost 70 percent of its shipments to meet commitments to customers. The failure of marketing, manufacturing, and logistics to communicate cost the company a substantial amount of money in increased freight costs.

Similar decisions need to be made in the design of the distribution system. The warehouse or distribution center location decision begins with the issues of how many warehouses and where to build them. For each warehouse, additional capital investment is required and added inventory

in the form safety stock must be carried. However, extra warehouses mean that distances from customers are reduced. The type of warehouse--manual to fully automated--also impacts the place and time utility a distribution system is able to deliver. Storage is just one aspect of the distribution system; movement is the second primary component. Two decisions dominate the movement of goods: modal choice and carrier selection. Modal choice focuses on which of the five modes of transportation to use--rail, motor carrier, air, water, and pipeline. The primary criteria for the modal choice decision include product characteristics, geographic reach, service requirements, and cost. While product characteristics and geographic reach determine modal feasibility, service and cost typically create a tradeoff situation--higher service modes are more expensive. Carrier selection identifies and selects the best set of transportation service providers based on a set of factors that emphasize service and cost comparisons. It is interesting to note that information technology plays a significant role in the design and management of today's distribution systems, impacting both storage and movement activities. Information substitutes for inventory throughout the distribution system, coordinates diverse activities, and integrates supply chain members. The following examples demonstrate that managers from information systems, logistics, marketing, and strategic planning all need to be involved in the design and management of distribution systems that are capable of delivering profitable customer takeaway via place and time utilities.

- Wal-Mart achieved much of its success by providing consistently low prices on a large variety of products that were almost always available on the shelf when customers wanted to buy them. To do this, Wal-Mart combined cross-docking, a private trucking fleet, an internal satellite system, and information system linkages among all of its stores and its major suppliers. This approach allowed Wal-Mart to know exactly how much product was available on the shelf as well as when and how much to order from suppliers. Consolidating orders allows bulk purchases and truckload shipping to the distribution center where the

majority of goods are cross docked to meet individual store needs and shipped again via full truck loads. This aggressive integration of information and distribution systems has allowed Wal-Mart to achieve outstanding efficiency with much more frequent delivery to retail store locations than its competitors. The bottom line is that Wal-Mart keeps its distribution costs low and its on-shelf stock levels high.

- National Semiconductor analyzed its competitive position and found that its market share was decreasing, largely as a result of long lead times and poor on-time delivery. This finding led managers to review the distribution system design and make some radical changes. The existing distribution system relied on regional distribution centers and local ground transportation. The revised system called for consolidating the operations of several regional warehouses into a single distribution center located in Singapore and shipping products directly to customers via air freight. While transportation costs increased, total logistics costs decreased. More importantly, lead times were cut in half and market share expanded rapidly [Henkoff, 1994].

The bottom line is that more information sharing, greater understanding, and increased coordination across functional boundaries is needed to achieve the efficiencies needed for profitability while using time and place utilities to help customers perform well along their critical success factors.

DISCUSSION AND SUGGESTIONS FOR FUTURE RESEARCH

We attempt to clarify the domain of profitable customer takeaway in the supply chain. Additionally, we define value creation in terms of real customer capabilities. Finally, we suggest that firms can help their customers succeed by integrating functional and inter-organizational activities that create and deliver valued capabilities.

The value of the profitable customer takeaway

concept and the associated notion of customer success is that they provide a powerful strategy-guiding influence that can help define the firm's organizational culture along more integrated, cross-functional lines. Indeed, when the competitive focus is on delivering high levels of profitable customer takeaway, customer requirements always remain in focus without obscuring the need for operational excellence. Profitable customer takeaway thus helps firms determine what to do as well as how to do it--two of the core tasks of an effective strategy. Achieving alignment between customer success factors and the firm's distinctive capabilities almost always enhances firm performance and long-term customer success. Future research would do well to examine empirically the relationships between customer success factors and the firm's distinctive capabilities in the supply chain. In addition, research directed at assessing the relationship between the firm's capabilities and the customers' customer success factors would be potentially highly rewarding for researchers of the supply chain.

Profitable customer takeaway also presents a rationale and a framework for motivating greater intrafirm and supply chain integration. The previous discussion of utility creation illustrated that the value-added contribution of most activities, even those that have long been viewed as falling under the domain of a particular function, can be enhanced through greater cross-functional integration and cooperation. This integrative theme holds true for almost all activities and processes that possess the distinctive value-added ability needed to truly deliver profitable customer takeaway. Consequently, we urge researchers to evaluate the effects of functional integration and cooperation on profitable customer takeaway. The results of such research would be of value to supply chain researchers and practitioners alike.

REFERENCES

- Anderson, James C. and James A. Narus (1999), *Business Market Management*, UpperSaddle River, NJ: Prentice Hall.
- Blackwell, Roger D. (1997), *From Mind to Market: Reinventing the Retail Supply Chain*, New York: Harper Collins Publishers.
- Drucker, Peter F. (1994), "The Theory of the Business," *Harvard Business Review*, 73, (5), 95-104.
- Fawcett, S. E. and S. A. Fawcett (1995), "The Firm as a Value-Added System: Integrating Logistics, Operations, and Purchasing," *International Journal of Physical Distribution and Logistics Management*, 25, (3), 24-42.
- Henkoff, R. (1994), "Delivering the Goods," *Fortune*, November 28, 64-78.
- Ohmae, Kenichi (1988), "Getting Back to Strategy," *Harvard Business Review*, 66, (6), 149-156.
- Patterson, Paul G., Lester W. Johnson and Richard A. Spreng (1997), "Modeling the Determinants of Customer Satisfaction for Business-to-Business Professional Services," *Journal of the Academy of Marketing Science*, 25, (Winter), 4-17.
- Prahalad, C. K. and Gary Hamel (1990), "The Core Competence of the Corporation," *Harvard Business Review*, (May-June), 79-91.
- Quinn, James Brian and Frederick G. Hilmer (1994), "Strategic Outsourcing," *Sloan Management Review*, 35, (Summer), 43-55.
- Scully, J. I. and S. E. Fawcett (1994), "International Procurement Strategies: Opportunities and Challenges for the Small Firm," *Production and Inventory Management Journal*, 35, (2), 39-46.
- Stalk, G., P. Evans and L. E. Schulman (1992), "Competing on Capabilities: The New Rules of Corporate Strategy," *Harvard Business Review*, 70, (2), 57-69.
- Yarbrough, J. F. (1996), "The Best Sales Forces," *Sales and Marketing Management*, (November), 46.

Send correspondence regarding this article to:

Stanley E. Fawcett
 Marriott School of Management
 Brigham Young University
 Provo, UT 84602 USA

CUSTOMER COMMENT CARDS IN THE SERVICE SECTOR: AN EMPIRICAL INVESTIGATION OF SCOPE AND FORMAT

**Frederick B. Kraft, Wichita State University
Charles L. Martin, Wichita State University**

The editors regret that in Volume 10, the first line of this article title was omitted. We profusely apologize for this error.



