Understanding the Structure of Consumers' Satisfaction Evaluations of Service Delivery

Jagdip Singh
Case Western Reserve University

This paper examines the current state of the literature concerning the structure (i.e., conceptualization, taxonomy, and operationalization) of consumers' satisfaction evaluations of a service delivery. In particular, relevant research in medical sociology, community health and marketing is reviewed. In so doing, advances as well as gaps in our current understanding of satisfaction evaluations are delineated. Drawing upon the services marketing, social psychology, and organizational theory literatures, we propose a hypothesis for the multiobject conceptualization of the satisfaction construct. This fills a major gap in previous research. Using data collected from four metropolitan areas, this hypothesis is subjected to empirical verification. Competing hypotheses are explicitly considered. The proposed conceptualization is also tested for convergent, discriminant, and nomological validity. Overall, the results support a multidimensional-multiobject model of the satisfaction construct. However, the objects are the major source of variation in satisfaction evaluations, accounting for about 61% of the total variance. In addition, object-based evaluations yield evidence of convergent, discriminant, and nomological validity. Implications for research into the structure of satisfaction evaluations for health care delivery in particular, and for multidyadic services in general, are discussed.

INTRODUCTION

The understanding of the process (i.e., consequences and antecedents) and the structure (i.e., content and dimensions) of consumer satisfaction appears to be a central concern for several constituencies. Researchers treat satisfaction as a key variable in models of consumer behavior (e.g., Howard and Sheth 1969); practitioners regard customer satisfaction as the focal point for designing successful marketing strategies (Dixon 1989); and public policy officials/consumerist agencies recognize satisfaction levels as barometers of consumer welfare (Czepiel and Rosenberg 1977).

Despite its importance, much previous research in marketing has tended to focus mainly on satisfaction processes, paying little attention to its structure. As such, Oliver and DeSarbo (1988, p. 495) observe that, "more so than others in related disciplines, consumer researchers have advanced and tested the processes underlying satisfaction, placing less emphasis on (its) content." This lack of emphasis is surprising because problems due to poor conceptualization and operationalization of the satisfaction construct have been well documented in the marketing literature (e.g., MSI/NSF sponsored conference on the conceptualization and measurement of satisfaction; Hunt 1977). Specifically, it has been noted that the understanding of the structure of the satisfaction construct is critical not only for measurement/tracking purposes (e.g., for public policy officials and practitioners) but also for providing greater insights into satisfaction processes. Consequently, Westbrook and Oliver (1981, p. 94) suggest that the inadequate development of the satisfaction construct "hinder(s) the interpretation and synthesis of (satisfaction) research findings."

The purpose of this paper is to contribute to the understanding of the structure of consumer satisfaction data. Specifically, several alternative structures, based on a priori hypotheses, are investigated in the context of consumers' evaluations of their health care service. The choice of health care service was guided by three factors. First, in sharp contrast to marketing, the medical sociology and community health literatures have made significant advances in mapping out the structure of consumers' satisfaction with health care delivery (Zyzanski, Hulka and Cassel 1974; Locker and Dunt 1978; Ware, Davies-Avery and Stewart 1978; Mangelsdorff 1979; Penchansky and Thomas 1981). This offers an opportunity for marketing scholars to exploit the "pioneering advantage" of these literatures and integrate...
with the research in marketing. Some attempts in this direction have recently appeared (Swan and Carroll 1980; Westbrook and Oliver 1981; Pascoe 1983). However, because these studies have not been published in the main marketing journals, it is likely that many researchers are unaware of these efforts.

Second, issues concerning the structure of satisfaction are especially critical for services. In contrast to products, services are usually intangible (e.g., medical care), often involve the customer in the "production" process (e.g., college education) and result from direct provider-client interaction (e.g., physician-patient) (Shostack 1977; Love-lock 1983). This increases the complexity of consumers' satisfaction evaluations (Parasuraman, Zeithaml and Berry 1985; Surprenant and Solomon 1987). Little empirical work, however, has been done to understand such evaluations for services.

Third, the measurement of consumers' satisfaction with health care delivery is itself a critically important issue. In particular, the turbulent environment of the 1980s is marked by regulation (e.g., enforcement of diagnosis-related groups), aggressive cost-cutting strategies, and competitiveness among health care providers. Although this environment may well hold, if not decelerate, spiraling medical costs, it is less clear at this time if the price of controlling cost increases will be paid in the coin of poor quality and lower patient satisfaction. Thus, consumer satisfaction concerns have become central issues for researchers, health care administrators, public policy officials, and consumers themselves (e.g., see Blendon and Altman 1984).

Below, first a review of the (patient) satisfaction (PS) construct in the medical sociology and community health literatures is presented, and areas where the marketing literature can contribute to or gain from this body of research are identified. Following this review, the paper draws upon the services marketing, social psychology, and organization theory literatures to build the case for incorporating multiple objects in satisfaction evaluations. Thus, unlike most previous research, this study explicitly hypothesizes that a service delivery system is not a homogeneous entity. Rather, it is composed of multiple objects or constituencies, such as the physician, the hospital, and the insurance provider in the case of health care service. In the third major section, the structure of consumer satisfaction data is examined empirically. In so doing, several competing models based on a priori hypotheses are explicitly examined. This investigation utilizes restricted factor analysis estimated by the method of Weighted Least Squares (WLS) via LISREL VII. This method represents a significant advantage in that it does not assume that Likert-type data have interval properties. Instead, it views such data as only ordinal in nature. As such, this approach departs from other conventional methods (e.g., ML estimation) in that it yields asymptotically correct standard errors of parameter estimates and $\chi^2$ goodness-of-fit statistics. In addition, tests for the convergent, discriminant, and nomological validity are also performed. Finally, the findings and limitations of the study are discussed, and implications for future research into the satisfaction construct in general and satisfaction with health care delivery in particular are outlined.

A REVIEW OF THE MEDICAL SOCIOLOGY AND COMMUNITY HEALTH LITERATURE

Several excellent reviews have appeared in the medical sociology and community health literatures (Locker and Dunt 1978; Ware et al. 1978; 1983). In addition, the journal Evaluation and Program Planning devoted two full issues (1983, vol. 6) to the research involving the PS construct. These sources facilitate a focused and succinct discussion of the conceptual, taxonomical, and operationalization issues for the satisfaction construct. Readers are directed to appropriate references for detailed analysis and review.

Conceptual Issues

Initial attempts at conceptualizing the (patient) satisfaction (PS) construct are identified with Hulka and her associates (Hulka, Zyzanski, Cassel and Thompson 1970; Zyzanski, Hulka and Cassel 1974). They defined "satisfaction" as the consumer's "attitude toward physicians and medical care" (Hulka et al. 1970, p. 430). More specifically, they hypothesized a composite index of an individual's evaluative judgments concerning the quality of medical care received from physicians, nurses and other relevant sources to represent his/her level of "satisfaction." Within the community health literature, this conceptual definition has been widely accepted (Linder-Pelz 1982; Hines et al. 1977; Doyle and Ware 1977; Ware et al. 1978).

However, refinements in this definition have been advanced from at least three perspectives. The first perspective notes that an episode (or situation) is a major source of variation in satisfaction evaluations. This does not imply that consumers cannot or do not make global satisfaction judgments based on an aggregate of health care episodes. Rather, satisfaction studies should distinguish between micro (i.e., episode-specific) and macro (i.e., global) evaluations because empirical studies show that consumers do evaluate these aspects differently (Shore and Franks 1986). In addition, several researchers have argued that it may be more useful to conceptualize satisfaction as an individual's evaluation of the quality of care in a specific medical-care situation, not just as a global attitude aggregated across episodes. This argument is exemplified by a recent paper by Shore and Franks (1986). These researchers note that individual patient-physician encounters are "the basic unit of medical care" (p. 580) and, therefore, assessing satisfaction for "individual encounters may contribute to a fuller understanding of the nature of physician-patient relationship." Similar arguments have been advanced by Inui and Carter (1985) and Pascoe (1983). Consistent with this view, marketing researchers utilize an episode based conceptualization for the satisfaction construct (Oliver 1981; Westbrook and Oliver 1981). In fact, Westbrook and Oliver (p. 95) have criticized the PS literature in general, and Hulka et al.'s conceptual definition in particular, because it taps "generalized satisfaction with the overall domain of primary health care rather than evaluation of a specific experience."

Second, the conceptualization of satisfaction as an cognitive evaluation (i.e., attribute based) versus an emotional state (i.e., global feeling) has generated some debate. Mar-
Several attempts for developing a taxonomy of satisfaction attributes have been reported in the literature. Hulka, et al. (1970) proposed one of the earliest satisfaction taxonomy. While these researchers do not provide the precise approach utilized, it appears that a two-step strategy was adopted. First, the literature was reviewed to determine “the appropriate content areas around which statements concerning attitudes towards physicians and medical care should be developed” (p. 430). Next, these content areas were subjectively classified into three dimensions: (a) professional competence, (b) personal qualities, and (c) cost/convenience.

For an alternative taxonomy, Ware, Davies-Avery and Stewart (1978) utilized a content analysis of the published literature and open-ended responses from about 700 consumers. They found support for an eight dimensional structure: (a) art of care, i.e., provider characteristics relating to the manner (or mode) of health care delivery (e.g., friendliness); (b) technical quality of care, i.e., professional competence; (c) accessibility/convenience, i.e., time and effort required to obtain medical care; (d) finances, i.e., cost/benefits of medical care; (e) physical environment, i.e., characteristics of the care environment (e.g., comfort, cleanliness, etc.); (f) availability, i.e., number of health care providers available; (g) continuity of care, i.e., regularity of the health care provider; and (h) efficacy/outcomes of care, i.e., improvement in health status attributable to medical care obtained. In the same article, however, Ware et al. reported that empirical support for the various dimensions was mixed. Specifically, it was noted that the “most well documented” empirical dimensions were physician conduct (i.e., art and quality of care) and accessibility/convenience. Furthermore, empirical attempts to distinguish between the art of care and quality of care were “much less convincing.”

More recently, Smith, Bloom and Davis (1986) have sought to integrate previous taxonomical attempts in the community health literature with research in the marketing literature (e.g., Swan and Combs 1976). They have proposed a tripartite taxonomy for the satisfaction construct, consisting of: (a) expressive, implying notions of art of care: (b) instrumental, including factors such as quality of care, efficacy of treatment, and continuity of care; and (c) access/cost, i.e., notions of accessibility, convenience and cost.

For several reasons, Smith et al.’s integrative taxonomy has merit. First, note that Smith et al.’s tripartite taxonomy subsumes the one proposed by Hulka et al. (1970). Second, two of three dimensions in Smith et al.’s taxonomy (i.e., instrumental and access/cost) were found by Ware et al. (1978) to be stable and consistent in several different data. Third, evidence in support for the significance and distinction between the expressive and instrumental dimensions is forthcoming from the medical sociology literature. In particular, Ben-Sira (1976, 1980) utilized the social interaction theory to posit that the affective (or expressive) component of the physician’s behavior toward the patient is a major factor in the assessment of the instrumental component of the physician’s care. As such, Smith et al.’s taxonomy achieves considerable parsimony by classifying a wide range of attributes into three theoretically useful and empirically stable dimensions. For the preceding reasons, it appears desirable to utilize Smith et al.’s tripartite taxonomy in future research.

Operationalization Issues

Several attempts have been made to develop operational measures for the satisfaction construct. Hulka, et al. (1970) proposed one of the earliest operationalizations: a 42 item scale for assessing three dimensions of satisfaction. Of the
three dimensions, only the personal quality dimension had an alternate forms reliability greater than 0.7. For the professional competence and the access dimensions, the reliability was 0.63 and 0.43 respectively. When different samples were utilized, similar levels of reliability were found by Hulka and her associates.

Some researchers have proposed operational measures for one or more specific dimensions of the satisfaction. For instance, Aday and Anderson (1975) focused on cost-convenience and provider characteristics. Likewise, Bice and his colleagues (e.g., Bice and Kalimo 1971) developed measures for availability/convenience, physician conduct and utilization tendencies. More recently, Penchansky and Thomas (1981) provided an operational scale to measure the “access” component of satisfaction evaluations. These researchers posited that the access component itself is multidimensional and provided empirical support for six distinct dimensions, namely, availability, accessibility, accommodation, affordability, and acceptability.

Unlike the preceding attempts, Ware and his associates (e.g., Ware and Snyder 1975) conducted an extensive methodological study to operationalize most, if not all, potential dimensions of the PS construct. Based on a review of most existing scales, open-ended responses from consumers, and multisample analysis, an eighty item operationalization was proposed (PSQ, Form I). This measure was conceptually designed to assess twenty dimensions of the satisfaction concept. Empirically, however, Ware and Snyder found support for four basic factors, namely, physician conduct, availability of service, continuity/convenience of care, and access to care. Further analysis revealed that these four basic factors in turn represent just two second-order dimensions, pertaining to physician conduct-quality and accessibility/availability (Roberts et al. 1983).

Within the medical sociology and community health literatures, each of the preceding operationalizations have been utilized with some regularity. For instance, Tucker and Tucker (1985) report a study in which Hulka et al.’s operational measure was incorporated. In a study of primary care selection, Sullivan (1984) utilized Ware et al.’s operational measure. Finally, an example of a study that used the Penchansky and Thomas scale is the research by Tucker and Tucker (1985).

Despite this acceptance, caution in the use of the preceding measures is warranted for several reasons. First, these operationalizations measure satisfaction at the macro (i.e., global) level. To the extent that researchers are interested in micro (i.e., episode-specific) satisfaction, these measures cannot be directly utilized. However, some attempts to reword the preceding measures for micro level ratings have been reported (Weinberger et al. 1981).

Second, most of the preceding studies do not pay adequate attention to indirect versus direct measurement of satisfaction (cf. Pascoe 1983). A direct measure asks the respondent to articulate how s/he would evaluate her/his own health care service. An example of this item is, “The doctor has relieved my worries about my illness.” By contrast, the indirect measure does not explicitly focus on respondent’s own health care service. Rather, it leaves the focus ambiguous. An example of this item is, “Doctors always do their best to keep the patient from worrying.” In this sense, many of the preceding operationalizations are mixed; they contain direct as well as indirect items. After critically reviewing this issue, Pascoe (1983, p. 190) observes that there is enough “logical and empirical evidence” to conclude that “there is not a strong relationship between direct and indirect measures of satisfaction.” For this reason, the inattention to direct/indirect measures in previous operationalizations is problematic.

Third, some researchers have cautioned against the use of the preceding operational measures for certain testing situations (e.g., Hausknecht 1988). Specifically, Hausknecht observes that a summary measure of consumers’ satisfaction with health care delivery may be desirable in some contexts, such as when testing antecedents and/or consequences of satisfaction. In such situations, researchers have tended to sum up the satisfaction ratings for individual attributes to arrive at overall satisfaction. This approach is arbitrary because it makes two assumptions: (a) consumers utilize a compensatory mechanism (i.e., summing up) in making overall satisfaction judgments, and (b) consumers give equal weight to individual attributes. Because these assumptions are untested, the use of attribute-based satisfaction evaluations (e.g., Ware et al.’s PSQ) may be less useful despite their richer insight. As an alternative, Hausknecht recommends an overall episode-specific measure of satisfaction. Consequently, it appears useful (in some situations) to include an overall measure of satisfaction with a health care service experience in addition to attribute-based evaluations. This allows a direct assessment of how the individual aspects contribute to overall consumer satisfaction; thus, avoiding the preceding arbitrary assumptions.

Finally, it has been noted that the satisfaction construct may have been underconceptualized (Pascoe 1983; Singh 1988). For instance, Pascoe (1983) notes that the conceptualization of PS may have been myopic in its focus on provider (i.e., physician) dimensions. As a result, non-provider dimensions (e.g., ancillary services) or other aspects of health care (cf. Singh 1988) are inadequately considered in satisfaction evaluations. This issue is discussed in greater detail below.

INTEGRATING MULTIPLE OBJECTS IN CONSUMERS’ SATISFACTION EVALUATIONS

Although the medical sociology and community health literatures provide insights into the structure of satisfaction evaluations (i.e., questions along the lines of “what is consumer satisfaction?”), relatively less attention has been directed at issues such as “what is the consumer satisfied with?” in terms of identifying the object (e.g., physician, nurse) of such evaluations. As a result, in most operationalizations the satisfaction object is treated unsystematically. Consider, for instance, the eighty item PSQ measure proposed by Ware et al. Of the 80 items in PSQ, 57 items pertain to satisfaction with a doctor, 4 utilize the hospital as the object of reference, another 4 concern medical insurance, and the remaining 15 are general satisfaction items.

This lack of systematic attention to objects is surprising because research in services marketing, social psychology,
and organizational behavior indicates that objects are an important source of variation in consumers' evaluations. This is especially valid for "customer contact" (i.e., consumer involvement is needed in service delivery; Chase and Tansik 1983), "multiobject" (i.e., service involves interactions with multiple service personnel; Solomon et al. 1985), and "loosely coupled" services (i.e., the different organizations and/or departments involved in service delivery are loosely related; Weick 1976). These features manifest in several commonly used services, such as the health care service, restaurants, airlines, banks, and education. Figure 1 displays these features for health care service. Below, three streams of research which directly bear on the case for systematically incorporating multiple objects in satisfaction evaluations are summarized. Specifically, pertinent findings from the marketing (i.e., "service encounters and scripts"), organization behavior (i.e., "boundary spanning roles") and management science (i.e., "customer contact model") are discussed.

**Service Encounters and Scripts**

The marketing literature recognizes that health care delivery is a relatively "pure" form of service. Specifically, this literature posits three prototypical attributes of services: (a) intangibility, that is services are relatively more intangible than products, (b) inseparability of production and consumption, that is services tend to be produced and consumed more simultaneously than products, and (c) customer participation, that is consumers actively participate in service delivery. Although different services possess more or less of these attributes, the health care service mostly evi-
dences all of the preceding attributes (Lovelock 1983; Sil-
pakit and Fisk 1985). In particular, for such “pure” services, the
quality of service delivery rests to a large degree on the
way in which the provider−consumer interaction (i.e., ser-
vice encounter) proceeds and, consequently, it is unpre-
dictable a priori (Solomon et al. 1985). Furthermore, the service
delivery is likely to be heterogeneous across interaction
episodes.

In terms of satisfaction evaluations, this performance
ambiguity coupled with service attributes (e.g., intangibility
etc.) causes difficulties for consumers in evaluating the ser-
vice received. For this reason, marketing researchers posit
that consumers utilize the quality of the provider−consumer
interaction itself as the basis for evaluating the service
received. For instance, Solomon et al. (1985, p. 100) observe
that in the case of pure services, “customer satisfaction and
repeat patronage may be determined solely by the quality of
personal encounter.” Likewise, Ben-Sira (1976, 1980) has
supported this perspective on the basis of social interaction
theory. Consequently, the service encounter is a focal point
for considering satisfaction evaluations.

It is the case with many services that the service encoun-
ters are “multiobject”—that is, they involve multiple ser-
vie personnel (i.e., objects) in distinct and separate dyadic
encounters. Consider the case of a medical service that
requires hospitalization. The consumer most likely interacts
(to more or less extent) with at least three distinct service
providers: (a) the physician for diagnosis and treatment, (b)
hospital and its staff2 (e.g., nurses) for care and testing, and
(c) insurance personnel for payment of services. Clear dif-
ferences exist in the types of personnel involved and the
goals sought in the various encounters. As a result of these
differences, it is logical to expect that consumers may eval-
uate these encounters differently. Thus, it is plausible that,
in a given health care experience, the consumer is very
satisfied with his/her physician but not satisfied with the
hospital staff. At the same time, the same consumer may be
very dissatisfied with the insurance personnel.

Theoretical developments in social psychology concern-
ing script theory provide further evidence in support of mul-
tiobject evaluations. Smith and Houston (1983) have argued
that consumers utilize and/or possess “service scripts”. These
scripts contain information about the role-set (i.e.,
one’s own expected behavior and the expected behavior of
service provider) in dyadic service encounters. Further-
more, this line of research posits that consumers possess
different service scripts for different types of service en-
counters (e.g., visit to dentist, obtaining auto-repair).

For the health care example (see Figure 1), it follows that
consumers may possess a physician script, a hospital staff
script, and an insurance script. Evidence supporting these
differences emerges from the medical sociology literature
(Parsons 1975; Waitzkin 1985; Ben-Sira 1980). For in-
stance, in the case of physician−patient encounter, Buller
and Buller (1987) observe that the “institutionalized roles”
for the physician and patient favor greater “power, author-
ty, professional detachment and status” in the physician’s
role. Clearly, this is less valid for patient interactions with
hospital staff and insurance personnel. These disparate ser-
vie scripts suggest that consumers may have different ex-
pecations (i.e., role-set) in interacting with different ob-
jects in a given service system. This lends support to the
argument that objects may be an important source of varia-
tion in satisfaction evaluations.

**Boundary Spanning Roles**

In contrast to marketing’s focus on consumer-provider
encounters, the organizational researchers have documented
and analyzed features of service organizations (Mills, Chase
and Marguiles 1983). In this research, the boundary span-
ning roles (BSR) are usually the focal point of study (Bowen
and Schneider 1988). Organizational theorists have noted
that BSR personnel represent organizations by (a)
acquiring and/or disposing services/products/resources, (b)
maintaining the image of the organization, and (c) sustain-
ning or improving the organization’s legitimacy (cf. Aldrich
1979). BSR personnel are especially important in services
because they often interact with the consumer in the cre-
ation of the service and because consumers tend to rely on
their behaviors in forming service evaluations (Bowen and
Schneider 1988).

In multiobject service systems (e.g., health care; Figure
1), several BSR personnel (e.g., physician, nurse) are in-
volved in service delivery. More importantly, these BSR
personnel typically belong to organizations that are func-
tionally and physically either completely decoupled or only
loosely coupled with each other (Weick 1976). In Figure 1,
note that the insurance companies are usually completely
decoupled from hospitals and physicians. Furthermore,
physicians represent a professional organization that is only
loosely coupled with the hospital’s administrative organiza-
tion. This is even more valid for private practitioners who
are “attached” to different hospitals merely for the use of
their facilities and staff. For these reasons physicians often
consider themselves outside of the hospital’s administrative
hierarchy and control.

Furthermore, these decoupled or loosely coupled organi-
zations often differ in their goals, objectives, and the image
they wish to present to their public. For instance, in terms of
their objectives, the insurance provider may be interested in
balancing the coverage with the premiums, the hospital in
increasing the bed occupancy rate, and the physician in the
physical health of the consumer. Such differences most like-
ly are reflected in BSR’s behaviors and expectations—i.e.,
BSR’s scripts of service encounters. Because these scripts
influence how the service encounter proceeds, it follows
that differences in BSR’s scripts (e.g., due to organizational
affiliation) would be important sources of variation in multi-
object service encounters. Consequently, utilizing object
based satisfaction evaluations may not only capture real
differences in service encounters, but may also be desirable
from an organization perspective (e.g., for diagnosis, BSR
training).

**Customer Contact Model**

Management scientists (e.g., Chase and Tansik 1983) and
service marketers (e.g., Langeard et al. 1981) have noted
that the extent of customer contact (i.e., high or low) is an
important factor in services organizational design and clas-
sification. In their customer contact model, Chase and Tan-
sik argue that service organizations which require high contact in service delivery most likely differ (from those that require low contact) in terms of design and operation due to efficiency and effectiveness considerations. Likewise, Silpakit and Fisk argue that differences in customer participation and contact underlie the various factors (e.g., see Love-lock 1983) utilized to classify services.

In a health care service, important differences in customer contact are evident for the multobject encounters. In a physician–patient interaction, a high level of customer contact can be anticipated. For hospital’s staff–patient service encounters, the extent of contact is likely to vary from low (e.g., with registration staff) to high (e.g., with nurses). By contrast, the interactions between the insurance personnel and the patient may not even be face-to-face. As a result, such interactions represent low customer contact. Because the extent of customer contact is a critical factor in services research (Chase and Tansik 1983; Langeard et al. 1981), it is reasonable to posit that encounters with wide variation in customer contact should not be lumped together. Instead, they should be treated as distinct and heterogeneous evaluations. This argues for incorporating multiple objects in consumers’ satisfaction evaluations.

The preceding evidence emerging from three different literatures appears consistent and compelling enough to hypothesize that consumer satisfaction can be accurately understood as a collection of multiple satisfactions with various objects that constitute the service system. In particular, for the case of health care service, three distinct objects are posited: the physicians, the hospital, and the insurance providers. Likewise, multiple objects can be identified for other services that evidence the features noted in Figure 1.

Note, however, that the preceding is merely a hypothesis for the multobject structure of satisfaction evaluations. Furthermore, since it is based only on conceptual and theoretical arguments, it is not yet known if this hypothesis is a valid representation of how consumers actually evaluate the service. In addition, it is less clear if objects or dimensions (or both) are the major source of variation in satisfaction evaluations.

Thus, two competing hypotheses are explicitly considered so as to ascertain the substantive usefulness of the proposed hypothesis. The first competing hypothesis pertains to a multidimensional structure for satisfaction evaluations. In this hypothesis, the objects of satisfaction (e.g., physician, insurance provider) do not provide much information about consumers’ evaluations. Instead, attribute dimensions are sufficient to explain the structure of satisfaction evaluations. Based on the preceding review, a tripartite structure is considered, with expressive, instrumental, and access/cost as the hypothesized dimensions. Note, that this multidimensional view of satisfaction is implied in most previous research in medical sociology and community health.

The second competing hypothesis posits a multobject-multidimensional structure for satisfaction evaluations. According to this hypothesis, neither objects nor dimensions are sufficient in and of themselves to explain variations in consumers’ satisfaction evaluations. Instead, both are important factors. Note, that this hypothesis represents a full model (i.e., it contains objects and dimensions). Therefore, it is obvious that this model will be statistically superior to the preceding models. In terms of substantive and parsimony criteria, however, the full model may not be more desirable. For instance, it is possible that the full model fails to explain significantly more variance in satisfaction evaluations (i.e., after adjusting for the number of additional parameters estimated) than the competing models. Likewise, if the addition of objects (or dimensions) yields only incremental substantive understanding of satisfaction evaluations, the full model would be less attractive. For these reasons, substantive and parsimony criteria were utilized for evaluating the competing models, in addition to statistical criteria.

The explicit consideration of such competing hypotheses is desirable because it provides a more reasonable basis for evaluating the proposed hypothesis than the traditional null hypothesis of no effects. Most likely, empirical analysis is necessary to sort through these hypotheses. An initial study was designed to address these issues in the context of health care service.

THE STUDY

The study was designed to shed empirical light on the structure of the satisfaction construct in the context of health care delivery. An ideal test would have required rewording current satisfaction scales (e.g., Ware et al.’s eighty items) to obtain separate evaluations for the physician, hospital, and the insurance company. For Ware et al.’s scale, this implies in all about 240 items (eighty items times three objects). This was considered too unwieldy (in terms of questionnaire length) and premature given the current state of the literature. Instead, it was decided to opt for an elementary measurement for the individual dimensions and objects. The measurement was elementary in the sense that a small (e.g., compared with Ware’s PSQ) but representative set of items was utilized to measure the individual constructs. Specifically, six items to measure each dimension/object (in all eighteen items) were selected. It was felt that this approach was appropriate to provide initial insights into focal questions and guide future research (e.g., whether to emphasize dimensions or objects).

Pascoe (1983, p. 188) has noted that satisfaction evaluations would be unambiguous if the consumption context was “explicit and consistent.” In accord with this, a recent episode health care experience was explicitly defined as the unit of analysis for the present study. The use of a specific episode is consistent with Shore and Franks (1986) and Inui and Carter (1985). In addition, Pascoe (1983) provides empirical evidence supporting the argument that consumers can differentiate between (dis)satisfaction stemming from specific experiences and “global” health care. Respondents were asked to think about their recent experience so as to control for degrading effects of recall from memory.

Four different metropolitan areas were selected for conducting the study: Cleveland (Ohio), Jacksonville (Florida), Omaha (Nebraska) and Salt Lake City (Utah). In many previous studies, researchers have tended to sample from hospital lists without controlling for the metropolitan–rural mix of patients. Because patients who come from rural
areas may face different circumstances from those in metropolitan areas, it was considered desirable to control for this factor. For this reason the study identified metropolitan areas as the geographical domain for sampling patients. Although individual households within these areas were randomly selected, the areas themselves were not. Rather, a systematic procedure was developed in order to select these areas. This procedure was based on obtaining (a) geographical dispersion across the U.S., and (b) variation in the number of physicians per capita. The first condition controlled for an over-representation of a particular region within the US. The second condition controlled for restriction in range for the variation due to the extent of physician competition. While somewhat crude, the number of physicians per capita (computed from U.S. Census data) provided some estimate of this competition. Unfortunately, this data was not available for individual metropolitan areas. Thus, States were stratified into groups on the basis of physicians per capita, and the individual metropolitan areas were then selected to satisfy the first condition. The number of non-federal physicians per 1000 people in the four states selected were as follows: Nebraska = 1.55; Ohio = 1.69; Utah = 1.71; and Florida = 2.09.

Data Collection

Consumer self-reports through a mail survey was selected as the method for data collection. Self-reports are appropriate because satisfaction is defined as a perceptual construct. In addition, while mail surveys are susceptible to response-set bias (e.g., social desirability, cognitive consistency), Pascoe (1983, p. 195), after reviewing the effects of such artifacts, has observed “that while such effects do operate, patients’ self-reports are not substantially biased by these artifacts.” The population of interest was defined as the households who (a) had medical insurance coverage, since evaluations of insurance provider were important for this study, and (b) had a recent medical care experience. Sampling frames for such population are not easily available, however (except for hospital lists which have other limitations as noted above). For this reason, an alternative procedure was adopted in which a random sample of households were asked to preselect themselves if they can recall a recent hospital experience. Randomly selected households obtained from a mailing house for each of the four cities were mailed an identical questionnaire packet. In all, 1,500 questionnaires were mailed (375 for each city). Reminder cards (two) were used to encourage participation. About 10% of the surveys were returned because of nondelivery (e.g., due to moves, wrong addresses).

The number of responses received for analysis as were follows: Cleveland = 154, Jacksonville = 103, Omaha = 133, and Salt Lake City = 140. Differences in mean values for the various constructs (see measures below) were examined across the four sampling areas. The null hypotheses for no significant differences could not be rejected for the satisfaction measures (all F-values < 1.80, p > 0.05). Thus, it seems appropriate to pool data for analysis. However, because evaluations of insurance providers are essential for our research, only the 432 respondents with insurance coverage were retained for analysis. Furthermore, cases with one or more missing values were deleted, leaving 367 usable responses.

True response rates cannot be estimated precisely since this involves computing the proportion: (households who responded/ households who have insurance coverage and had a recent hospital experience). As noted above, the denominator of this term is an elusive number. However, secondary data provides an estimate for the proportion of households in the US who are without medical insurance coverage. Specifically, Kiesler and Morton (1988) estimate that about 17% of the US population is uninsured. Assuming that the selected geographical areas reflect trends in the overall US population, and adjusting for a 10% nondelivery rate, a lower bound estimate for the overall response rate in this study is 39%, with a usable rate of 34%. These are lower bound estimates because the response rate is not adjusted for the proportion of the households who did not have a recent hospital experience. Unfortunately, secondary data is not available to impute this term.

Responses in the range obtained here are not uncommon in PS research (Pascoe 1983). For instance, in research designs similar to the one utilized in this study, Harris (1978) has noted that response rates can be as low as 18%. While less than perfect, response rates in this study are about two-fold greater than the lower bound reported by Harris. This suggests that the degrading effects due to nonresponse may not be excessive.

Nevertheless, an attempt was made to investigate nonresponse bias. Although, an analysis of nonrespondents was not possible because of the promised anonymity of responses, the data was examined for possible nonresponse bias via wave analysis (Armstrong and Overton 1977). Wave analysis was performed by classifying responses into three categories based on the postmark date. Responses to the various measures of study (see details below) were examined for significant differences in the three waves. The null hypothesis for no systematic differences could not be rejected (F values < 2; p > 0.05). The demographic profile of the respondents (for the pooled data) is in the Appendix. Because the survey instructions asked respondents to complete the questionnaire only if they could briefly describe a health care experience (i.e., requiring hospital visit) which they remember clearly, and involves either themselves or a member of their family for whom they make decisions (e.g., children), the respondents’ demographic profile can not be directly compared with the census profile of the corresponding geographical areas. Females constitute 84% of the respondents indicating that women play a major role in health care decisions. This is consistent with previous research (e.g., The Keckley Report 1986). The median age group is 36–40 years, but respondents range in age groups from 21–25 years (12% of sample) to over 60 years (16%), with no particular age group dominating the data. Fully 78% of the respondents are married, and 54% of the sample has had some education beyond high school. In terms of income, the median level is the range $20,001 to $30,000. Over 88% of the sample earns less than $50,000. In terms of race, however, our sample is somewhat skewed, since 95% of the sample is white. Although this may affect the representativeness of the sample, note that the focus of this study is relationships among variables, and not mean values.
per se. For this reason the skew in some demographic characteristics should not pose a severe limitation.

**Measures**

The satisfaction items were measured utilizing a six point “very satisfied—very dissatisfied” scale. The specific items were based on Ware et al.’s research. As indicated above, the items were selected so as to provide elementary measurement of each of the three dimensions. Most of Ware et al.’s items are applicable for physicians. Thus, items were modified/reworded to be relevant for hospital staff and medical insurance providers. In addition, all items were worded so as to yield a micro (i.e., episode specific) measure of satisfaction. Table 1 lists these items. Note, there are six items for each dimension, and for each object. The Cronbach’s α for the satisfaction measures utilized (see results below) were 0.83, 0.85, and 0.92 corresponding to the physician, hospital and insurance provider, respectively.

In addition, four “overall” satisfaction items were included. These items did not tap respondent’s global satisfaction across several prior experiences. Rather, they were intended to provide an overall assessment of satisfaction in the specific recent experience stemming from (a) physician, (b) hospital staff, (c) insurance provider, and (d) overall health care service. For instance, in the case of physician, the respondents were asked, “Overall, with the care provided by my doctor, I felt . . . .” These overall items are in accord with Hausknecht’s (1988) suggestion that satisfaction should be assessed as an overall response to a specific episode. They also facilitate the investigation of convergent and discriminant validity of satisfaction items (discussed below).

Finally, additional items were included in order to investigate nomological validity of satisfaction items. These items tap behavioral intentions to switch physicians, hospitals, and/or insurance services in the future. Several studies provide empirical evidence for the nomological relationship between satisfaction and self-reports of switching intentions (DiMatteo, Prince and Taranta 1979; Needle 1975). The behavioral intentions were measured by a six item scale (two each for physician, hospital, and insurance provider). Table 1 lists these items as well. Response were obtained on a six point, “very unlikely—very likely” scale. The Cronbach’s α for the switching intentions measures were 0.82, 0.79, and 0.65 corresponding to the physician, hospital and insurance provider, respectively.

**Method of Analysis**

The proposed hypothesis for the multiobject satisfaction evaluations was evaluated by three procedures. First, four alternative conceptual models were proposed as possible representations of data. These models correspond to: (a) the null model, $M_0$, positing that there are no underlying factors, (b) the multidimensional model, $M_1$, hypothesizing that the three satisfaction dimensions are sufficient to explain intercorrelations among items (see Figure 2), (c) the multiobject model, $M_2$, positing that the three satisfaction objects are sufficient (see Figure 3), and (d) the multidimensional multiobject model, $M_3$, which proposes that both dimensions and objects are necessary to represent satisfaction evaluations (see Figure 4).

These models were then evaluated by estimating the competing models with LISREL VII (Jöreskog and Sörbom 1988); therefore, the Figures are depicted in the conventions of latent variable analysis. Specifically, circles represent latent constructs, boxes refer to observables, arrows connecting boxes and circles are measurement relations, and arrows without origins are error terms. The use of LISREL is desirable because of several reasons. First, it allows estimation of restricted factor analysis with a priori constraints on loadings. These constraints correspond to the specific model being estimated. Second, it provides a systematic basis for evaluating competing measurement models. This evaluation is based on various goodness-of-fit statistics. For each model, LISREL VII provides a $\chi^2$ statistic (and associated degrees of freedom) corresponding to the test for the null hypothesis that the intercorrelations reproduced by the hypothesized model equal the observed correlation matrix. Also, a goodness-of-fit index (GFI), an adjusted goodness-of-fit index (AGFI; this adjusts the GFI for degrees of freedom), and root mean square residual (RMR) for the model estimated are provided. Unlike the $\chi^2$ statistic, the GFI, AGFI, and RMR are less sensitive to the sample size. Following Bagozzi and Yi (1988), GFI and AGFI values exceeding 0.90, and low RMR values (typically less than 0.05) are indicative of models that are “good” representations of data. Third, two competing models (say $M_1$ and $M_2$) can be tested by utilizing the Bentler and Bonnet’s (1980) normed fit index (NFI). This index provides a measure for the improvement in fit between models $M_2$ and $M_1$ as compared with the null model ($M_0$). NFI values exceeding 0.90 are indicative of adequate fits. Although competing models can also be tested by computing a $\chi^2$ difference statistic (i.e., difference between the $\chi^2$ of $M_2$ and $M_1$), this test is not technically appropriate here because the competing models are not properly nested within each other (Anderson and Gerbing 1988). However, for pedagogical reasons, this statistic is computed as well. Finally, additional assumptions about the data are not required in LISREL VII. In particular, for data obtained on a Likert-scale, it is not necessary to assume that such data have interval properties. Instead, LISREL VII treats such data as categorical with underlying continuous latent distribution(s) by computing polychoric correlations and analyzing by the method of Weighted Least Squares (WLS). This represents a significant advantage because methods that assume that Likert-type data have interval properties can “lead to greatly distorted parameter estimates and incorrect $\chi^2$ goodness-of-fit measures and standard errors” (Jöreskog and Sörbom 1988, p. 192). This troublesome situation is avoided in LISREL VII by the use of WLS method.

Of the competing models, a particular model(s) is selected as the most appropriate, based on statistical, parsimony, and substantive criteria. The statistical criterion stems from the overall goodness of fit statistics (e.g., $\chi^2$, GFI). One or more models are accepted as plausible representations of data if each model provides acceptable overall statistics (e.g., nonsignificant $\chi^2$, NFI > 0.90, etc.). By contrast, the parsimony criterion argues for a simpler model over a complex model if both provide acceptable statistical
TABLE 1
Operational Measures for the Constructs of Study

<table>
<thead>
<tr>
<th>Object</th>
<th>Dimension</th>
<th>Mean</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physician</td>
<td>Expressive</td>
<td>4.85</td>
<td>With my doctor's personal concern for me, I feel . . (#1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.25)</td>
<td></td>
</tr>
<tr>
<td>Physician</td>
<td>Expressive</td>
<td>4.85</td>
<td>With my doctor's willingness to explain reasons for medical treatment, I feel . . (#2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.21)</td>
<td></td>
</tr>
<tr>
<td>Physician</td>
<td>Instrumental</td>
<td>5.03</td>
<td>With my doctor's medical abilities, I feel . . (#3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.04)</td>
<td></td>
</tr>
<tr>
<td>Physician</td>
<td>Access/Cost</td>
<td>4.21</td>
<td>*With my doctor's choice of hospitals, I feel . . (#4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.44)</td>
<td></td>
</tr>
<tr>
<td>Physician</td>
<td>Access/Cost</td>
<td>3.91</td>
<td>With my doctor's ability to see me on time, I feel . . (#5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.46)</td>
<td></td>
</tr>
<tr>
<td>Hospital</td>
<td>Expressive</td>
<td>4.34</td>
<td>With my hospital's personal concern for me, I feel . . (#7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.18)</td>
<td></td>
</tr>
<tr>
<td>Hospital</td>
<td>Instrumental</td>
<td>4.71</td>
<td>With my hospital's willingness to explain its procedures, I felt . . (#8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.04)</td>
<td></td>
</tr>
<tr>
<td>Hospital</td>
<td>Instrumental</td>
<td>4.71</td>
<td>With my hospital's medical capabilities, I felt . . (#9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.12)</td>
<td></td>
</tr>
<tr>
<td>Hospital</td>
<td>Access/Cost</td>
<td>3.27</td>
<td>With my hospital's costs, I felt . . (#11)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.44)</td>
<td></td>
</tr>
<tr>
<td>Insurance</td>
<td>Expressive</td>
<td>3.64</td>
<td>With my insurance provider's personal concern for me, I felt . . (#13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.41)</td>
<td></td>
</tr>
<tr>
<td>Insurance</td>
<td>Instrumental</td>
<td>3.91</td>
<td>With my insurance provider's willingness to explain its procedures, I felt . . (#14)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.39)</td>
<td></td>
</tr>
<tr>
<td>Insurance</td>
<td>Access/cost</td>
<td>3.73</td>
<td>*With my insurance provider's location, I felt . . (#12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.44)</td>
<td></td>
</tr>
</tbody>
</table>

Switching Intentions

<table>
<thead>
<tr>
<th>Object</th>
<th>Mean</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physician</td>
<td>1.64</td>
<td>In regard to my physician, I will . .</td>
</tr>
<tr>
<td></td>
<td>(1.15)</td>
<td>1. continue to use my current doctor (reverse scored).</td>
</tr>
<tr>
<td></td>
<td>2.21</td>
<td>2. look for other doctors.</td>
</tr>
<tr>
<td></td>
<td>(1.46)</td>
<td></td>
</tr>
<tr>
<td>Hospital</td>
<td>1.80</td>
<td>In regard to my hospital, I will . .</td>
</tr>
<tr>
<td></td>
<td>(1.14)</td>
<td>1. continue to use my current hospital (reverse scored).</td>
</tr>
<tr>
<td></td>
<td>2.28</td>
<td>2. look for other hospitals.</td>
</tr>
<tr>
<td></td>
<td>(1.36)</td>
<td></td>
</tr>
<tr>
<td>Insurance</td>
<td>1.70</td>
<td>In regard to my insurance provider, I will . .</td>
</tr>
<tr>
<td></td>
<td>(1.07)</td>
<td>1. continue to use my current policy (reverse scored).</td>
</tr>
<tr>
<td></td>
<td>2.67</td>
<td>2. look for other policy.</td>
</tr>
<tr>
<td></td>
<td>(1.65)</td>
<td></td>
</tr>
</tbody>
</table>

*This item was deleted on the basis of initial analysis.

a Standard deviation in parenthesis. For all scales, the means are based on a six point scale. For satisfaction items, the scale points were as follows: 1 = Very Dissatisfied, 2 = Dissatisfied, 3 = Somewhat Dissatisfied, 4 = Somewhat Satisfied, 5 = Satisfied, and 6 = Very Satisfied. For behavioral intention items, the scale points were as follows: 1 = Very Unlikely, 2 = Unlikely, 3 = Somewhat Unlikely, 4 = Somewhat Likely, 5 = Likely, and 6 = Very Likely.
b The item number in parenthesis at the end of each item is utilized to refer corresponding items in Figures 2, 3 and 4.

fits. A model is simpler if it posits fewer latent constructs and/or measurement relationships. Finally, the substantive evaluation is based on examining the amount of variation attributable to objects, dimensions, and error in satisfaction data. Utilizing the full model (i.e., M3), the variance in each satisfaction item is partitioned into the three components by computing the squares of standardized loadings (i.e., $\lambda^2$). Note that $\lambda^2$ provides an estimate for the variance shared...
FIGURE 2
The Hypothesized Structure for the Multi-dimensional Model of Patient Satisfaction

FIGURE 3
The Hypothesized Structure for the Multi-object Model of Patient Satisfaction
between the item and the corresponding factor. By so decomposing each item's variance, it can be ascertained if objects and/or dimensions are the major sources of variation in satisfaction evaluations.

Second, the convergent and discriminant validity of the multi-object satisfaction evaluations was examined. To accomplish this, a multitrait-multimethod (MTMM) matrix was analyzed by the method of hierarchically nested covariance structure models (Widaman 1985). In the MTMM for this study, there are three different traits (i.e., physician, hospital, and insurance satisfaction), each of which is assessed by two methods, namely the overall item and the multi-item ratings (see measures section, above). Because these methods are not maximally dissimilar, this matrix is designated as a pseudo-MTMM. There is precedence for such an approach (e.g., Cadotte, Woodruff and Jenkins 1987). Also note that the overall items were not utilized in the first procedure. The posited model for the

FIGURE 4
The Hypothesized Structure for the Multi-dimensional Multi-object Model of Patient Satisfaction
UNDERSTANDING THE STRUCTURE OF CONSUMERS’ SATISFACTION EVALUATIONS OF SERVICE DELIVERY

The Model for the Analysis of Pseudo-Multitrait-Multimethod Consumer Satisfaction Data

P-MTMM (see Figure 5) as well as other nested models were analyzed using LISREL VII by the method of WLS because of the reasons outlined above. Widaman (1985) suggests that the evidence for convergent and discriminant validity can be obtained by comparing the change in $\chi^2$ between the unrestricted model (i.e., in which the methods and traits are allowed to correlate freely; see Figure 5) and an appropriate nested model. In particular, for discriminant validity, the appropriate nested model is the one in which all of the correlations among the traits are fixed to unity (model 2C in Widaman 1985). For this nested model, if the change in $\chi^2$ is statistically significant (based on the difference in degrees of freedom), discriminant validity among the traits is tenable. Likewise, for convergent validity, the appropriate nested model is the one in which the loadings for the traits are set to zero and the inter-correlations among the traits are also set to zero (model 1C in Widaman 1985). Such a nested model attempts to fit the data with correlated method factors only. As such, if the change in $\chi^2$ is statistically significant for this model, the traits are likely to possess convergent validity.

Finally, the nomological validity of the multiobject satisfaction evaluations was investigated. A structural model was posited (see Figure 6) in which the individual object satisfactions (e.g., physician) were modeled as antecedents of behavioral intentions to switch the physician, hospital, or the insurance provider. For the multiobject hypothesis to be valid, structural coefficients must be higher (in absolute value) within the same object (e.g., physician satisfaction intentions to switch physician) than across objects (e.g., physician satisfaction intentions to switch hospital). Consistent with the previous models, the nomological network of Figure 6 was estimated by the method of WLS using LISREL VII. Note, the structural coefficients from this procedure are corrected for measurement error, and thus represent a reasonable basis to examine nomological validity within objects, and a lack thereof across objects.

RESULTS

Because the satisfaction items were specifically developed for this study, an initial analysis was performed to determine the quality of items. This analysis was based on
exploratory factor analysis and examination of inter-item correlations. Of the 367 usable responses, 150 were randomly selected for this analysis. The selection of a subset of the sample was preferred because it would not compromise the confirmatory nature of the study. This analysis revealed that 3 of the 18 items (see Table 1) were inconsistent with the remaining items. These items had low correlations with other items in its set and had poor loadings and/or cross-loadings in factor analysis. This loading pattern was evident even when six factors were extracted. For these reasons, the three items (indicated in Table 1) were deleted from further analysis. Univariate statistics (i.e., means and standard deviations) are also included in Table 1. The following analysis is based on the remaining 15 items.

**Confirmatory Analysis of Alternative Models**

*Overall Model Fit.* Results for the alternative conceptual models of satisfaction are shown in Table 2. Based on the overall goodness-of-fit statistics, the $M_3$ model (i.e., Figure 4) yields satisfactory fit statistics (i.e., $\chi^2 = 103$, $df = 70$, $p = 0.006$, GFI = 0.99, AGFI = 0.98, and RMR = 0.04), indicating that the reproduced correlations nearly equal observed correlations in this model. Although the $\chi^2$ is statistically significant, problems with the overall $\chi^2$ test even in moderately large samples (e.g., $n > 200$; in our case $n = 367$) are well documented (Bentler and Bonnet 1980; Bagozzi and Yi 1988). For this reason, other indicators of fit (e.g., GFI, AGFI, RMR) are given greater prominence.

Comparisons of $M_3$ with $M_2$ and $M_1$ (see last two rows in Table 2) indicate that $M_3$ yields significant improvement in NFI relative to the $M_1$ model (value = 0.25), and comparatively smaller, though still significant, improvement relative to the $M_2$ model (value = 0.05). Furthermore, $\chi^2$ differences for $M_3$ with all other competing models are sig-
TABLE 2

Tests for the Alternative Models of the Consumer Satisfaction Construct

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
<th>Model Tests¹</th>
<th>Model Comparisons²</th>
</tr>
</thead>
<tbody>
<tr>
<td>M₀</td>
<td>Null</td>
<td>χ² 8469 df 120 GFI 0.25 AGFI 0.25 RMR 0.40</td>
<td>d d d</td>
</tr>
<tr>
<td>M₁</td>
<td>Figure 2</td>
<td>χ² 2210 df 87 GFI 0.80 AGFI 0.73 RMR 0.20</td>
<td>M₀-M₁ 6259 df 33 NFI 0.74</td>
</tr>
<tr>
<td>M₂</td>
<td>Figure 3</td>
<td>χ² 527 df 87 GFI 0.95 AGFI 0.93 RMR 0.09</td>
<td>M₀-M₂ 7942 df 33 NFI 0.94</td>
</tr>
<tr>
<td>M₃</td>
<td>Figure 4</td>
<td>χ² 103 df 70 GFI 0.99 AGFI 0.98 RMR 0.04</td>
<td>M₁-M₃ 2107 df 17 NFI 0.99</td>
</tr>
</tbody>
</table>

¹All models were estimated by the method of Weighted Least Squares (WLS) using LISREL VII. The matrix of polychoric correlations was input to this method along with the asymptotic covariance matrix of the correlation matrix.
²All chi-square differences are significant at p = 0.01. The difference in the chi-square values of the models under comparison.
³NFI is the normed fit index based on Bentler and Bonnet (1980).
⁴For this model, a single Heywood condition was obtained for the theta delta parameter for item #15. This parameter was fixed at a very small value (0.001) and the model reestimated.

significant at p = 0.01. Other indicators of fit (e.g., AGFI, RMR) provide additional evidence that neither M₂ (note high RMR and large χ²) nor M₁ (note poorer on all statistics) are acceptable representation of data. This suggests that M₃ and only M₃ is acceptable as a plausible representation of consumer satisfaction evaluations.

Sources of Variation. Recall that the M₃ model posits that satisfaction is a multidimensional multiobject construct (Figure 3). Substantively, however, it is less clear if the dimensions or the objects (or both) are the major source of variance in satisfaction ratings. This is an important issue because it provides guidelines for the further development of the satisfaction construct. For instance, if the dimensions are the major source of variation, then future operationalizations must explicitly focus on dimensions, and the loss of information by ignoring objects may be negligible. By contrast, if both dimensions and objects share equal variance, arguments for ignoring either dimensions or objects in future research would be less persuasive.

Using the LISREL results for the M₃ model, Table 3 depicts the individual item variance partitioned into three parts: due to (a) dimension, (b) object, and (c) unique and random error. The corresponding WLS parameter estimates are in Table 4. Evidence here suggests that the objects are the major source of variance. First, note in Table 4 that estimates for object factor loadings are consistent and uniformly higher than the corresponding loadings for dimensions. Second, Table 3 reveals that object variance ranges from 31% to 77%, with an average contribution of 61%. By contrast, the dimensions produce isolated effects. For instance, the access/cost dimension (note: item pertains to costs) accounts for a significant portion of the variance for physician (item 6) and hospital (item 11) evaluations. The expressive dimension appears critical for the physician (variance contribution 11% and 14%) and the insurance provider (variance contribution of 14% and 10%). On the average, the dimensions account for only about 11% of the variance. Thus, the overall conclusion here is that the object evaluations constitute the major factors in consumers' satisfaction judgments. Next, we examine if the object evaluations achieve convergent, discriminant, and nomological validity.

Convergent and Discriminant Validity

Initially the unrestricted P-MTMM model of Figure 5 was estimated using the WLS procedure in LISREL VII. However, it was not possible to utilize the raw variables directly because problems were encountered in estimating the asymptotic covariance matrix of the polychoric and polyserial correlations. This mainly occurred due to the "small" sample size relative to the number of variables (Jöreskog and Sörbom 1988, p. 192; note the number of variables have increased from the measurement models of Figures 2–4). To address this problem, the raw indicants for each of

TABLE 3

Sources of Variation in Satisfaction Evaluations¹

<table>
<thead>
<tr>
<th>Item²</th>
<th>Objects</th>
<th>Dimensions</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physician Items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item #1</td>
<td>70</td>
<td>11</td>
<td>19</td>
</tr>
<tr>
<td>Item #2</td>
<td>74</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>Item #3</td>
<td>64</td>
<td>0</td>
<td>48</td>
</tr>
<tr>
<td>Item #4</td>
<td>49</td>
<td>3</td>
<td>38</td>
</tr>
<tr>
<td>Item #6</td>
<td>31</td>
<td>31</td>
<td>35</td>
</tr>
<tr>
<td>Hospital Items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item #7</td>
<td>77</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>Item #8</td>
<td>74</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>Item #9</td>
<td>70</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>Item #10</td>
<td>38</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>Item #11</td>
<td>42</td>
<td>32</td>
<td>61</td>
</tr>
<tr>
<td>Insurance Items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item #12</td>
<td>74</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>Item #13</td>
<td>77</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>Item #14</td>
<td>66</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Item #15</td>
<td>61</td>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td>Item #16</td>
<td>49</td>
<td>29</td>
<td>33</td>
</tr>
</tbody>
</table>

¹All values are in percentages.
²Corresponds to items listed in the order in Table 1.
³Computed as the square of the corresponding factor loading.
## Table 4
WLS Parameter Estimates for the Multidimensional Multiobject Model of Satisfaction Evaluations

<table>
<thead>
<tr>
<th>Objects</th>
<th>Physician</th>
<th>Hospital</th>
<th>Insurance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expressive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instrumental</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access/cost</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLS Factor Loadings&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Item #1</td>
<td>0.83</td>
<td>0.34</td>
</tr>
<tr>
<td>Item #2</td>
<td>0.86</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Item #3</td>
<td>0.80</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Item #4</td>
<td>0.70</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>Item #5</td>
<td>0.56</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td>Item #6</td>
<td>0.88</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Item #7</td>
<td>0.86</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Item #8</td>
<td>0.84</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Item #9</td>
<td>0.62</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Item #10</td>
<td>0.65</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>Item #11</td>
<td>0.86</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Item #12</td>
<td>0.88</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>Item #13</td>
<td>0.80</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>Item #14</td>
<td>0.78</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>Item #15</td>
<td>0.70</td>
<td>0.54</td>
<td></td>
</tr>
</tbody>
</table>

Goodness-of-Fit Measures
- Chi-square: 103
- Degrees of freedom: 70
- Goodness of Fit Index: 0.99
- Adjusted Goodness of Fit Index: 0.98
- Root Mean Square Residual: 0.04

<sup>a</sup>Corresponds to items listed in the order in Table 1.

## Table 5
WLS Parameter Estimates for the Model in Figure 5: Test for Discriminant and Convergent Validity

<table>
<thead>
<tr>
<th>Parameter&lt;sup&gt;a&lt;/sup&gt;</th>
<th>WLS Estimate</th>
<th>t-value</th>
<th>Parameter&lt;sup&gt;a&lt;/sup&gt;</th>
<th>WLS Estimate</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physician Satisfaction</td>
<td></td>
<td></td>
<td>Method 1 (Multi-item Ratings)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{1,1}$</td>
<td>0.70</td>
<td>2.88</td>
<td>$\lambda_{1,4}$</td>
<td>0.71</td>
<td>1.98</td>
</tr>
<tr>
<td>$\lambda_{2,1}$</td>
<td>0.13</td>
<td>0.51</td>
<td>$\lambda_{2,4}$</td>
<td>0.69</td>
<td>3.95</td>
</tr>
<tr>
<td>$\lambda_{1,1}$</td>
<td>0.62</td>
<td>2.71</td>
<td>$\lambda_{3,4}$</td>
<td>0.51</td>
<td>1.12</td>
</tr>
<tr>
<td>Hospital Satisfaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{3,2}$</td>
<td>0.70</td>
<td>2.92</td>
<td>$\lambda_{4,4}$</td>
<td>0.76</td>
<td>3.69</td>
</tr>
<tr>
<td>$\lambda_{4,2}$</td>
<td>0.47</td>
<td>1.48</td>
<td>$\lambda_{5,4}$</td>
<td>0.51</td>
<td>3.07</td>
</tr>
<tr>
<td>$\lambda_{8,2}$</td>
<td>0.59</td>
<td>2.65</td>
<td>$\lambda_{6,4}$</td>
<td>0.50</td>
<td>2.34</td>
</tr>
<tr>
<td>Insurance Satisfaction</td>
<td></td>
<td></td>
<td>Method 2 (Overall Ratings)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{3,3}$</td>
<td>0.64</td>
<td>2.92</td>
<td>$\lambda_{7,5}$</td>
<td>0.80</td>
<td>3.26</td>
</tr>
<tr>
<td>$\lambda_{6,3}$</td>
<td>0.72</td>
<td>2.43</td>
<td>$\lambda_{8,5}$</td>
<td>0.70</td>
<td>2.83</td>
</tr>
<tr>
<td>$\lambda_{9,3}$</td>
<td>0.79</td>
<td>3.04</td>
<td>$\lambda_{9,5}$</td>
<td>0.54</td>
<td>2.80</td>
</tr>
</tbody>
</table>

Intercorrelations: Satisfaction
- $\phi_{21}$: 0.31, $\phi_{31}$: 0.16, $\phi_{32}$: 0.24

Goodness-of-Fit Measures
- Chi-square: 3.46
- Degrees of freedom: 17
- p-value: 0.99
- Goodness of Fit Index: 0.99
- Adjusted Goodness of Fit Index: 0.99
- Root Mean Square Residual: 0.05

<sup>a</sup>All parameters correspond to symbols in Figure 5.
the satisfaction objects (e.g., physician) were split randomly into two groups to provide alternate indicators for each object. There is precedence for such an approach (Bagozzi 1980; Michaels, Day and Joachimsthaler 1987). Because of this, Figure 5 depicts two indicators for each object satisfaction. The WLS estimates for the model in Figure 5 are in Table 5.

Results in Table 5 suggest that the model in Figure 5 is a reasonable representation of the P-MTMM data ($\chi^2 = 3.46$, $df = 17$, $p = 0.99$, GFI = 0.99, AGFI = 0.99, RMR = 0.05). Note, however, that in estimating this model a Heywood condition was obtained because of the very large correlation between the two methods. This is not surprising because the methods represent only different types of items (i.e., multiitem versus overall) and do not constitute two distinct or maximally dissimilar methods. For this reason, the correlation between the two methods was constrained to 0.95.

As noted earlier, the test for discriminant validity among the three object satisfactions was based on a nested model which constrained the intercorrelations among the traits to unity. When compared to the model in Figure 5, the change in $\chi^2$ was 47.7, $df = 3$, $p < 0.001$. This resoundingly rejects the hypothesis that the traits are perfectly correlated, thus suggesting that the three object satisfactions possess discriminant validity. Note, also that the estimates of trait intercorrelations in Table 5 are not very high; specifically, ranging from 0.16 to 0.31. This further supports the discriminant validity of the multiobject satisfaction evaluations.

### Table 6

<table>
<thead>
<tr>
<th>Parameter</th>
<th>WLS Estimate</th>
<th>Standardized Estimate</th>
<th>$t$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Loadings: Physician Satisfaction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{1,1}$</td>
<td>1.00$^a$</td>
<td>0.60</td>
<td>—</td>
</tr>
<tr>
<td>$\lambda_{2,1}$</td>
<td>0.94</td>
<td>0.57</td>
<td>5.51</td>
</tr>
<tr>
<td><strong>Loadings: Hospital Satisfaction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{3,2}$</td>
<td>1.00$^a$</td>
<td>0.56</td>
<td>—</td>
</tr>
<tr>
<td>$\lambda_{4,2}$</td>
<td>0.92</td>
<td>0.52</td>
<td>6.54</td>
</tr>
<tr>
<td><strong>Loadings: Insurance Satisfaction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{5,3}$</td>
<td>1.00$^a$</td>
<td>0.51</td>
<td>—</td>
</tr>
<tr>
<td>$\lambda_{6,3}$</td>
<td>0.72</td>
<td>0.36</td>
<td>3.24</td>
</tr>
<tr>
<td><strong>Loadings: Physician Switching Intentions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{1,1}$</td>
<td>1.00$^a$</td>
<td>0.93</td>
<td>—</td>
</tr>
<tr>
<td>$\lambda_{2,1}$</td>
<td>0.94</td>
<td>0.88</td>
<td>13.82</td>
</tr>
<tr>
<td><strong>Loadings: Hospital Switching Intentions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{3,2}$</td>
<td>1.00$^a$</td>
<td>0.93</td>
<td>—</td>
</tr>
<tr>
<td>$\lambda_{4,2}$</td>
<td>0.85</td>
<td>0.79</td>
<td>14.17</td>
</tr>
<tr>
<td><strong>Loadings: Insurance Switching Intentions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{5,3}$</td>
<td>1.00$^a$</td>
<td>0.89</td>
<td>—</td>
</tr>
<tr>
<td>$\lambda_{6,3}$</td>
<td>0.84</td>
<td>0.75</td>
<td>9.13</td>
</tr>
<tr>
<td><strong>Intercorrelations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Phi_{12}$</td>
<td>—</td>
<td>0.67</td>
<td>5.05</td>
</tr>
<tr>
<td>$\Phi_{31}$</td>
<td>—</td>
<td>0.36</td>
<td>4.68</td>
</tr>
<tr>
<td>$\Phi_{32}$</td>
<td>—</td>
<td>0.52</td>
<td>6.20</td>
</tr>
<tr>
<td><strong>Structural Coefficients</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_{11}$</td>
<td>$-1.54$</td>
<td>$-1.00$</td>
<td>$-5.72$</td>
</tr>
<tr>
<td>$\gamma_{21}$</td>
<td>0.21</td>
<td>0.13</td>
<td>0.41</td>
</tr>
<tr>
<td>$\gamma_{31}$</td>
<td>$-0.35$</td>
<td>$-0.23$</td>
<td>—</td>
</tr>
<tr>
<td>$\gamma_{12}$</td>
<td>$-0.25$</td>
<td>$-0.15$</td>
<td>—</td>
</tr>
<tr>
<td>$\gamma_{22}$</td>
<td>$-1.68$</td>
<td>$-1.01$</td>
<td>$-8.53$</td>
</tr>
<tr>
<td>$\gamma_{32}$</td>
<td>$-0.44$</td>
<td>$-0.28$</td>
<td>—</td>
</tr>
<tr>
<td>$\gamma_{13}$</td>
<td>$-0.02$</td>
<td>$-0.01$</td>
<td>—</td>
</tr>
<tr>
<td>$\gamma_{23}$</td>
<td>$-0.27$</td>
<td>$-0.15$</td>
<td>—</td>
</tr>
<tr>
<td>$\gamma_{33}$</td>
<td>$-1.76$</td>
<td>$-1.00$</td>
<td>$-7.92$</td>
</tr>
<tr>
<td><strong>Goodness-of-Fit Measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-square</td>
<td>61.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>degrees of freedom</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goodness of Fit Index</td>
<td>0.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted Goodness of Fit Index</td>
<td>0.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Root Mean Square Residual</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$This parameter was fixed to 1.00 to fix the scale of measurement.
In order to test for convergent validity, a nested model was estimated which included method factors only. This produced a change in the $\chi^2$ of 136.48, df = 12, $p < 0.001$ (i.e., compared to the model in Figure 5). This suggests that intercorrelations among the variables cannot be attributed to method factors only. As such, a significant proportion of the covariation among the indicators is uniquely accounted for by the satisfaction objects. This strongly supports convergent validity of the multiobject satisfaction construct.

**Nomological Validity**

Figure 6 depicts the model tested for examining the nomological validity of the multiobject evaluations. This model is consistent with the preceding analysis in that the satisfaction measures were grouped into two alternative indicators for each object by grouping the raw variables. However, the inclusion of all of the structural coefficients created problems of instability and large standard errors. To address this, initially only within object structural coefficients were estimated (i.e., diagonals of the $\gamma$ matrix). Following this, three additional models were estimated by freeing up the across object structural coefficients separately for each object satisfaction. This allowed the comparison of respective within and across object coefficients, for each object. Results from these analysis are in Table 6. Note that the loadings and overall goodness-of-fit measures are reported for the initial model only (i.e., in which the within object coefficients were estimated).

The results in Table 6 suggest that the model in Figure 6 fits the data reasonably well ($\chi^2 = 61.12$, df = 50, $p = 0.13$, GFI = 0.98, AGFI = 0.98, and RMR = 0.15). In addition, all of the loading coefficients are significant (i.e., $t$-values $> 2.0$) indicating that the individual constructs are measured well. More importantly, the nomological validity criterion, that within-object coefficients be greater than across-object coefficients (in absolute value), is strongly supported. For instance, the standardized coefficient between satisfaction with physician and intentions to switch physician is $-1.00$. This is the within-object coefficient. In comparison, the standardized coefficients across object for physician satisfaction (e.g., physician satisfaction and intentions to switch hospitals) are only 0.13 and $-0.23$. Likewise, hospital and insurance satisfaction ratings produce similar results. On the average, the within-object coefficient is $-1.00$. By contrast, the average across-object coefficient is only $-0.12$. This supports the nomological validity of the multiobject satisfaction evaluations.

**DISCUSSION AND IMPLICATIONS**

The purpose of this paper has been three-fold: (a) to review the medical sociology and community health literature on the structure of satisfaction evaluations; (b) to develop a hypothesis, based on theoretical arguments drawn from the services marketing, social psychology, and organization theory literatures, for the multiobject conceptualization of the satisfaction construct; and (c) to examine this hypothesis empirically using consumers’ evaluations in the context of health care delivery. Below, the results of this study are discussed and specific implications for theory, service practitioners, and future research are identified. First, however, some limitations to this study are enumerated.

**Limitations**

Not unlike other cross-sectional survey studies, the findings of this research should be evaluated in light of certain limitations. The results are based on a mail survey of four metropolitan areas. Although these areas were systematically selected, the findings may have limited generalizability due to this geographic restriction. Further, the high proportion of white and female respondents is likely to result in restriction of range and argues for replication with other health care users. Note, however, that the focus of the study is relationships among variables rather than mean values per se. The response rates in this study were of the order of around 39%. Also, additional responses were lost because respondents had provided incomplete responses yielding a usable rate of 34%. While response rates of this order are not unusual in PS research (Pascoe 1983), they may affect the validity of the results. Finally, the elementary measurement of various dimensions and the initial nature of the study suggests future replications and validation.

**Discussion**

Marketing researchers have tended to focus on the satisfaction processes, paying inadequate attention to the structure of satisfaction evaluations. By contrast, the medical sociology and community health literatures have closely examined the structure of patient satisfaction. However, our review of these literatures revealed that, although advances have been made in studying "what is consumer satisfaction?" in terms of the theoretical approach, dimensional structure (of attribute evaluations), and operationalizations, relatively little attention has been given to "what is the consumer satisfied with?" in terms of explicitly identifying the object of satisfaction evaluations. This state of affairs leaves an impression that objects are secondary, if not inconsequential, factors in satisfaction data. By contrast, dimensions appear to be regarded as the primary source of variance.

Our study attempted to examine critically the preceding state of affairs. Theoretically, compelling arguments based on the notions of service encounters (services marketing), service scripts (social psychology), boundary spanning roles in loosely coupled organizations, and customer contact (organizational theory) were forthcoming to support the hypothesis for incorporating multiple objects in satisfaction evaluations. Empirically, this initial study provided clear and compelling insights into the structure of satisfaction evaluations. Although, of the competing models, the multidimensional-multiobject model is unequivocally supported as an acceptable representation of data, the partitioning of item variances indicated that the objects were the major source of variation in consumers' evaluations. By contrast, dimensions were secondary factors contributing on the average just one-fifth as much variance as objects. In addition,
dimensions produced isolated effects. Thus, it seems appa-
rent that the unsystematic treatment of objects in satisfaction
evaluations, as evident in previous research, is unfortunate
and should not be continued. Instead, researchers should
carefully delineate the various objects in a service system
(cf. Figure 1) and explicitly include them in satisfaction
measurements.

Furthermore, the findings of this study show that con-
sumers draw important distinctions between the various ob-
ject satisfactions and may arrive at an overall satisfaction
judgment that reflects evaluations for the individual objects.
This conclusion stems from the evidence for the internal
consistency, convergent, discriminant, and nomological va-
idity of multi-object satisfaction evaluations. Each object
evaluation evidenced a high level of internal consistency
(α > 0.80). Using a pseudo-MTMM, clear evidence of con-
vergent and discriminant validity was obtained (Table
5). Consistent with this, nomological validity analysis indi-
cated that, for each object, within-object correlations be-
tween satisfaction ratings and switching intentions were at
least three-fold higher than across-object correlations (Table
6). Thus, it is apparent that object-based evaluations carry
substantially significant and critically important information
about consumers’ satisfaction evaluations. Consequently, in
most previous research, the (unintentional) lack of systemat-
ic attention to objects may have had the undesirable effect of
the proverbial “throwing out the baby with the bath water.”

Note, however, that the findings of this study do not
suggest that the multiobject view is sufficient to understand
consumers’ evaluations. Rather, a multiobject, multidimen-
sional structure is consistent with data. Thus, our findings
should not be taken to imply that dimensions are inconse-
quential. Instead, it is appropriate to conclude that future
analysis focusing on either dimensions or objects would
necessarily provide an incomplete picture of how con-
sumers actually evaluate their service encounters. For a
complete picture, there appears to be sufficient evidence to
conclude that both dimensions and objects should be ex-
licitly incorporated in satisfaction data.

This may represent a natural evolution of the satisfac-
tion construct from a concept concerned with evaluation of ag-
gregate (i.e., ignoring object differences) satisfaction to a
more specific formulation that specifies what particular ob-
ject in the service system serves as the focus of consumers’
evaluations. Similar evolution in the conceptualization and
focus has occurred for other constructs in marketing re-
search. Such evolutionary developments have often facil-
itated deeper understanding of the antecedents, con-
sequences, and the structure of such constructs. For in-
stance, a related construct in marketing—salesperson job
satisfaction—has evidenced similar development. Early
studies in job satisfaction viewed the construct as a global,
attitudinal response to job (e.g., Hoppock 1935). In further
development, Churchill, Ford and Walker (1974) concep-
tualized (and operationalized) the job satisfaction construct
(INDSALES) as composed of several satisfactions with dif-
ferent aspects of the job, many of which were distinct ob-
jects in the salespersons’ “job-system” (e.g., fellow work-
ers, supervisor, top management, customers). The use of
multifacet satisfaction has revealed richer insights into such
phenomena as turnover (Futrell and Parasuraman 1984),
role stress (Fry et al. 1986), and performance (Futrell and
that studies that “fail to take these unique [multi-facet]
characteristics into account are not likely to provide a com-
plete and accurate picture of reality.” A similar call appears
justified for the consumer satisfaction construct.

At a more general level, this study directs attention to and
offers some guidelines for conceptualizing the structure of
consumers’ satisfaction evaluations. Westbrook and Oliver
(1981) have speculated that assessing satisfaction with the
various aspects of the product/service may be fruitful be-
cause of its richer insights. Along these lines, this study
suggests that the various aspects of satisfaction fall into two
major types of evaluations. The first pertains to attribute-
based satisfaction evaluations. Advances in the medical so-
ciometry and community health literature suggest that these
attribute evaluations can be parsimoniously represented by a
tripartite structure consisting of expressive, instrumental,
and access/cost dimensions. We suspect that this dimen-
sional structure may be applicable for other, if not most,
services (e.g., banking, restaurants, airline travel, hotels,
college education; cf. Chase and Tansik 1983).

The object-based evaluations appear as the second major
component of satisfaction evaluations. Theoretically, this
study has argued that, for services that involve multiple
interactions with different people (i.e., objects), satisfaction
ratings might evidence significant variability due to objects.
Examples of such services include (a) college education,
where consumers interact with professors, administra-
tors, and staff, (b) banking services, where tellers, loan
officers, and other people may be involved, and (c) airline
travel, where travel agents, airline crew/hostess, and other
ground staff may interact with the consumer in service de-
ivery. For such services, inclusion of objects in satisfaction
evaluations appears rewarding. A recent study by Crosby
and Stephens (1987) underscores this recommendation.
Crosby and Stephens examined antecedents of satisfaction
in the case of whole life insurance. They conceptualized
satisfaction as having three elements, namely, satisfaction
with the (a) contact person, (b) core service, and (c) insti-
tution. Results of this study show that, in addition to being
distinct (i.e., possessing discriminant validity), the three
satisfaction components had differential relationships with
the modeled antecedents. For instance, boundary personnel
interactions (modeled by personal contact and customer ser-
vice) appeared to mostly influence satisfaction with contact
person only. Likewise, satisfaction with the core service and
institution yielded unique and nonredundant relationships
with the antecedents.

In summary, we advocate that marketing researchers pay
greater attention to the structure of satisfaction evaluations.
Especially in the case of multiobject services, we recom-
pend that marketing researchers may wish to entertain a
multidimensional-multiobject view of satisfaction evalua-
tions. We believe that this view potentially affords a richer
and more accurate picture of the nature and structure of
consumers’ satisfaction evaluations. In addition, adoption
of this view is likely to yield better understanding of how
consumers actually arrive at their satisfaction judgments
concerning consumption experiences (cf. Crosby and Stephens 1987). That is, our understanding of the process of satisfaction would be more insightful once issues pertaining to the structure of the satisfaction evaluations are satisfactorily addressed.

Implications

Our study offers three key implications. First, it suggests that further development of the satisfaction construct is both desirable and necessary. Issues that warrant the serious attention of researchers include: (a) explicitly recognizing the distinct objects in the service system, (b) developing operational measures that assess the unique dimensions for each object, and (c) conducting further psychometric studies for understanding the structure of satisfaction ratings. As an initial study in this direction, our research yields evidence in support of three distinct objects, namely physicians, hospitals, and insurance providers. Future research may wish to reproduce these results by utilizing more comprehensive measures. Also, replications using especially male and black/hispanic populations would provide useful insights into the validity of the obtained findings. In addition, other objects in the health care system could be identified. Among these lines Woodside, Frey and Daly (1988) examined service encounters during hospital stay and advanced the notion of satisfaction with four different service objects: specifically (a) admission and registration, (b) nursing care, (c) food service, and (d) housekeeping. This further decomposes the hospital satisfaction into finer evaluations. Similar decomposition may be achieved for insurance provider (cf. Crosby and Stephens 1987) and physician encounters. Future research should empirically examine whether consumers indeed make such fine evaluations. Such developmental work is important because it will facilitate better insights into issues such as, how consumers evaluate the health care system, why patients are satisfied (or dissatisfied), and what precisely can be done to improve satisfaction levels.

Second, the proposed conceptualization offers new avenues for investigation. Much previous research has asked if consumers are satisfied with their medical care. The results of this study suggest the pursuit of a new line of research questions. Examples of such questions are: Are some consumers more satisfied with their physicians than with their insurance providers? With what object are patients most dissatisfied? Do the satisfaction levels for the various objects differ? Should insurance agents be more customer oriented? The current operationalizations cannot sort through these possibilities. By contrast, the proposed conceptualization for the satisfaction construct can address such questions precisely.

ACKNOWLEDGMENT

The author would like to acknowledge the assistance of Carilynn B. Drummond in data collection.

NOTES

1. The term “structure” is utilized to represent the nature, content, and dimensions of the satisfaction construct. In this sense, discussion of structure issues is synonymous with construct development questions. In contrast, the “process” issues relate to the underlying cognitive/affective mechanisms which lead to the formation of satisfaction judgments. Oliver and DeSarbo (1988) have recently reviewed the various theoretical approaches for understanding such processes. Although some early work has tended to blur these distinctions by viewing satisfaction as a derived construct, contemporary research draws clear distinction between the state (and hence the nature and structure of the construct) of consumer satisfaction and the processes of its formation (e.g., see Churchill and Surprent 1982).

2. We recognize that several different objects are usually present within a hospital–patient interaction, such as nurses, registration staff, technicians, etc. (e.g., see Woodside, Frey and Daly 1989). For the purposes of this initial study, rather than introduce this additional

**APPENDIX I**

Demographic Characteristics of the Sample

<table>
<thead>
<tr>
<th>Sex</th>
<th>Education</th>
<th>Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>16%</td>
<td>Grad School 8%</td>
</tr>
<tr>
<td>Females</td>
<td>84%</td>
<td>College 37%</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26 to 30</td>
<td>12%</td>
<td></td>
</tr>
<tr>
<td>36 to 40</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td>46 to 50</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>50 to 55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10,000</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>$10,000 to $20,000</td>
<td>16%</td>
<td></td>
</tr>
<tr>
<td>$20,000 to $30,000</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>$30,000 to $50,000</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>&lt;$70,000</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>$70,000 to $90,000</td>
<td>19%</td>
<td></td>
</tr>
<tr>
<td>$90,000</td>
<td>19%</td>
<td></td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>78%</td>
<td></td>
</tr>
<tr>
<td>Separated</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>5%</td>
<td></td>
</tr>
</tbody>
</table>

PSQ may reveal that the major source of dissatisfaction is the “art of care,” or in other words, the “mode” (cf. Ben-Sira 1980) in which the health care was delivered. It is clear that improving the mode of health care delivery would, in this hypothetical case, increase satisfaction. However, what is less clear is who should be the target for such changes? Should nurses be trained to be more pleasant? or Should physicians spend more time communicating with patients? Should insurance agents be more customer oriented? The current operationalizations cannot sort through these possibilities. By contrast, the proposed conceptualization for the satisfaction construct can address such questions precisely.
complexity, we utilized the simpler notion that the "hospital" object represents the hospital itself (e.g., physical facilities) as well as encounters with the various hospital staff (e.g., nurses). In the implications section, however, we discuss this issue in greater detail. Also, in subsequent discussion we use the terms hospital and hospital staff interchangeably.

REFERENCES


Doyle, B., and J. Ware. 1977. "Physician Conduct and Other Factors That Affect Consumer Satisfaction with Medical Care." *Journal of Medical Education* 52 (October): 793–801.


**ABOUT THE AUTHOR**

Jagdip Singh is an Associate Professor of Marketing at the Weatherhead School of Management, Case Western Reserve University. His current research interests include consumer dissatisfaction and complaint responses, stress/dissatisfaction in boundary role personnel, and measurement issues in marketing. Professor Singh has published in the *Journal of Marketing, Journal of Marketing Research, Journal of Consumer Affairs*, and the *Journal of the Academy of Marketing Science*, among others.