WILLIAM BOULDING, AJAY KALRA, RICHARD STAELIN, and VALARIE A. ZEITHAML*

Relying on a Bayesian-like framework, the authors develop a behavioral process model of perceived service quality. Perceptions of the dimensions of service quality are viewed to be a function of a customer's prior expectations of what will and what should transpire during a service encounter, as well as the customer's most recent contact with the service delivery system. These perceptions of quality dimensions form the basis for a person's overall quality perception, which in turn predicts the person's intended behaviors. The authors first test this model with data from a longitudinal laboratory experiment. Then they develop a method for estimating the model with one-time survey data, and reestimate the model using such data collected in a field study. Empirical findings from the two tests of the model indicate, among other things, that the two different types of expectations have opposing effects on perceptions of service quality and that service quality perceptions positively affect intended behaviors.

A Dynamic Process Model of Service Quality: From Expectations to Behavioral Intentions

In response to the growing importance of services in the worldwide economy and the recognition by goods firms of the need to compete on the service dimensions of the augmented product, several researchers have examined the problems of measuring and managing service quality (Bitner 1990; Bolton and Drew 1991a,b; Parasuraman, Berry, and Zeithaml 1990; Parasuraman, Zeithaml, and Berry 1985, 1988; Zeithaml, Berry, and Parasuraman 1991). In this article, we add to this literature by providing insights into both the process by which customers form judgments of service quality and the way these judgments affect subsequent behavior. Specifically, we propose and estimate a process model of service quality that (1) traces the way customers form and update their perceptions of service quality and (2) identifies the consequences of these perceptions on individual-level behavioral intention variables that affect the strategic health of the firm.

*William Boulding is Associate Professor and Richard Staelin is the Edward and Rose Donnell Professor, Fuqua School of Business, Duke University. Ajay Kalra is Visiting Assistant Professor, Carnegie Mellon University. Valarie A. Zeithaml is Partner, Schmalensee-Zeithaml.

The authors thank seminar participants at MIT, Wharton, and Harvard, as well as France LeClerc, Julie Edell, and especially John Lynch for their helpful comments.

Our model development draws from the service quality, attitude, and customer satisfaction literatures. We follow the lead of the service quality literature and center our attention on modeling and measuring the cumulative construct of the overall quality level of a firm's service delivery system. We take note of the similarity between the construct "perceived service quality" from the service quality literature and the construct "attitude toward an object" from the attitude literature. This similarity helps us generate theoretical predictions in our model of service quality. We also draw from the satisfaction literature, though we make explicit the distinction between this literature, which emphasizes consumers' perceptions of a specific transaction, and the service quality literature, which emphasizes cumulative perceptions.[1]

At the core of our model is the assumption that individuals' current perceptions of the service quality of a firm just after a service contact are a blend of (1) their prior expectations of what *will* and what *should* transpire during the contact and (2) the actual delivered service during the service encounter. Further, we acknowledge

[1]Readers should not confuse the customer satisfaction measure discussed in the popular press and measured by many corporations with the satisfaction measure used in most academic satisfaction studies. The former is usually a cumulative concept whereas the latter is transaction specific. We discuss this difference subsequently.

Journal of Marketing Research
Vol. XXX (February 1993), 7–27

that consumers update their expectations whenever they receive relevant information about the service through such means as word-of-mouth, company communications, and contact with the firm's or the competitors' service delivery system.

We test our model with data from two different studies. The first was a laboratory study involving multiple service encounters within the setting of staying at a hotel. Two different prior expectations and the delivered service were manipulated. With these longitudinal data, we use standard experimental analysis to test our basic hypotheses. We then specify a formal structural model representing our conceptualization of the service quality process. Using the same experimental data, we simultaneously test our basic hypotheses and the specification of our structural equations.

The second study enabled us to increase the generalizability of our results by examining the service quality process for a different service by using a different research method (a field study). In the laboratory study we were able to control (and thus measure) the objective aspects of the delivered service, but in our field study we did not obtain any objective measures of the actual dimensions of the service encounter for each individual. In addition, we obtained measures of expectations and perceptions at only one point in time. Such data are common in the area of service quality where (1) customers normally are polled once to ascertain their expectations and perceptions and (2) actual service is not measured, partly because obtaining objective measures is difficult and partly because the actual service delivered normally varies from person to person (and server to server). Consequently, we develop a method of analysis based on our structural process model that controls for (removes) all unobserved, individual-specific information affecting the customer's expectations and perceptions (the actual service being one such factor) while still allowing estimation of two key parameters of our process model. Such a technique should have broad applicability to service firms that want to measure the relative influences of the two different expectations *and* the delivered service (despite the fact that it is unmeasured) on the customer's perceptions of the firm's service quality.

In addition to postulating and testing a new dynamic model of expectations and perceptions, and providing an analytic approach for estimating major portions of this model with multiple-measures data obtained at only one point in time, we add to the service quality literature in several other ways. Though other researchers have postulated the existence of different expectations, our study is the first empirical demonstration of the joint influence of our two postulated expectations in a service quality setting. We also link the satisfaction and service quality literature by showing our dynamic model of service quality to be compatible with the currently accepted definition of transaction-specific satisfaction. Further, because the major current empirical paradigm for assessing service quality (the gaps model proposed by Parasuraman, Zeithaml, and Berry 1985) and our model are a subset of a more general model, we are able to estimate and test the validity of these alternative conceptualizations. Finally, ours is one of the first published field studies in which individual-level data are used to examine empirically the impact of consumers' perceptions of service quality on a set of intended behaviors of strategic interest to the firm.

In the following section we develop our structural model and generate hypotheses for empirical testing. We then estimate the parameters of this model with the two different datasets. We conclude with a discussion of our results.

MODEL DEVELOPMENT

Because our model has many of the same constructs as prior models of service quality and customer satisfaction/dissatisfaction (CS/D), we begin this section with a brief review of the dominant concepts of these two literatures. Expectations and perceptions play an important role in both literatures. In general, both literatures treat these constructs as static, at least for estimation purposes. Also, recent studies in both literatures have acknowledged the existence of multiple classes of expectations (Forbes, Tse, and Taylor 1986; Tse and Wilton 1988; Wilton and Nicosia 1986; Zeithaml, Berry, and Parasuraman 1991). Two main standards of expectations emerge. One standard represents the expectation as a *prediction* of future events (Gilly 1979; Gilly, Cron, and Barry 1983; Miller 1977; Prakash 1984; Swan and Trawick 1980). This is the standard typically used in the satisfaction literature. The other standard is a *normative* expectation of future events (Miller 1977; Prakash 1984; Swan and Trawick 1980), operationalized as either desired or ideal expectations. This is the standard typically used in the service quality literature (Parasuraman, Zeithaml, and Berry 1988).

Though these literatures use different expectation standards, expectations and perceptions in both literatures are usually linked via the disconfirmation of expectations paradigm (Oliver 1977, 1980). This paradigm holds that the predictions customers make in advance of consumption act as a standard against which customers measure the firm's performance (Bearden and Teel 1983; Churchill 1979; Day 1977; Woodruff, Cadotte, and Jenkins 1983). In the CS/D literature this paradigm states that the higher the expectation in relation to actual performance, the greater the degree of disconfirmation and the lower the satisfaction (Bearden and Teel 1983; Latour and Peat 1979; Swan and Trawick 1981; Tse and Wilton 1988). Expectations also play a contrast, or disconfirming, role in the gaps model of service quality (Parasuraman, Zeithaml, and Berry 1985). In this model, the consumer's perception of overall service quality results from a comparison between expectations and perceptions of the different components of service. With perceptions of service held fixed, the higher the expectations, the lower the perceived quality.

Our model also includes expectations and perceptions. However, it differs from the disconfirmation formulation in that we postulate that individuals' overall quality assessments, and thus behaviors, are affected only by their current perceptions of the service, and not their current expectations. These current perceptions, in turn, are the result of customers' two types of prior expectations of the service and the most recent service encounter.

In developing our conceptualization, we organize our discussion around three processes: (1) the process by which customers form and update their expectations, (2) the process by which customers develop perceptions of the quality of specific aspects of the service delivery system as well as an overall assessment of the firm's service quality, and (3) the relationship between perceptions of overall service quality and intended behaviors. After describing each of these processes, we provide a summary of the model and its testable implications.

The Process That Generates Expectations

Customer expectations are pretrial beliefs about a product or service (Olson and Dover 1979). In the absence of any information, prior expectations of service will be completely diffuse. In reality, however, customers have many sources of information that lead to expectations about upcoming service encounters with a particular company. These sources include prior exposure to the service, word of mouth, expert opinion, publicity, and communications controlled by the company (e.g., advertising, personal selling, and price), as well as prior exposure to competitive services (Zeithaml, Berry, and Parasuraman 1991).

Following the example of recent work suggesting the importance of multiple expectation standards, we postulate two different classes of expectations. Consistent with the expectations-as-predictions standard often used in the CS/D literature, we propose that customers form expectations about what *will* happen in their next service encounter with a firm. We refer to these expectations as *will* expectations. We also propose that customers form expectations about what *should* happen in their next service encounter, that is, the service customers feel they appropriately deserve. This normative expectation, hereafter referred to as a *should* expectation, is close in spirit to the "what ought to happen" expectation proposed by Tse and Wilton (1988). We distinguish this *should* standard from the ideal, or desired, standard frequently used in the service quality literature (Zeithaml, Berry, and Parasuraman 1991). What customers think *should* happen may change as a result of what they have been told to expect by the service provider, as well as what the consumer views as reasonable and feasible on the basis of being told of a competitor's service or experiencing the firm's or the competitor's service. In contrast, the consumer's *ideal* expectation—what a consumer wants in an ideal sense—may be unrelated to what is reasonable/feasible and/or what the service provider tells the customer to expect. Moreover, because *ideal* expectations represent enduring wants and needs that remain unaffected by the full range of marketing and competitive factors postulated to affect the *should* expectation, we believe *ideal* expectations are much more stable over time than consumer expectations of what should occur.

We start our discussion by noting that expectations and perceptions can change over time. Also, as becomes clearer subsequently, we acknowledge that there are J unique dimensions of service quality for each of these constructs. Finally, we note that our approach is to first specify general functional relationships for the process that generates these expectations and perceptions. After testing these general relationships, we specify and test explicit functional forms. These explicit equations enable us to gain additional insights into the process as well as develop an approach for estimating the parameters of our model with cross-sectional data.

More formally, let WE_{ijt} be consumer i's *will* expectation for the j^{th} dimension of a service just after experiencing a service contact at time t; DS^*_{ijt} be the j^{th} component of the service delivered to person i at time t (as captured by factors such as the number of thank you's, the waiting time, etc., and where the * notation indicates a transaction-specific construct as opposed to a cumulative construct); and $\mathbf{X}_{it}$ be a vector of information variables other than the service contact influencing the person's *will* expectations of the service prior to a new service contact. We acknowledge that a person's *will* expectations just before a new service contact can differ from the expectations held just after the prior service contact because of the information $\mathbf{X}_{it}$ that enters the system between service encounters. However, in our subsequent empirical work we do not measure such information. Consequently, our approach is to control, but not explicitly model or test, for effects of external information.[2]

We hypothesize that a consumer's expectations of what will happen in subsequent contacts with the firm's service delivery system depend not only on the information obtained from the most recent service contact, but also on the expectations held just prior to the service contact. Such a formulation explicitly acknowledges that two different individuals may hold different expectations about future service contacts even when they experience an identical (in an objective sense) service encounter. This is equivalent to saying that biases are present and that these biases are due to prior expectations.

More formally, we specify the following functional relationship:

$$WE_{ijt} = f_1(WE_{ijt-1}, \mathbf{X}_{it}, DS^*_{ijt}). \quad (1)$$

Note that equation 1 assumes expectations are influenced by the actual encounter (DS^*_{ijt} in our notation) ver-

[2]For a more explicit statement of how these $\mathbf{X}$ variables might influence expectations, see Boulding et al. (1992).

sus the consumer's perceptions of the actual encounter. We acknowledge that the consumer's preceptions of the *particular* service encounter may, in fact, be used to update expectations. However, if we denote this perception as PS^*_{ijt}, we not only believe $PS^*_{ijt} = g(DS^*_{ijt})$, but also that DS^*_{ijt} is a very good proxy for PS^*_{ijt}, that is, there is a strong positive relationship between the two constructs. Because our empirical work has no direct measure of PS^*_{ijt}, we integrate out this unobserved variable, which leads us to use DS^*_{ijt} instead of PS^*_{ijt} in specifying the functional relationship given by equation 1.

In making predictions about the effects of delivered service and prior expectations on a consumer's updated expectations, we believe a Bayesian-like updating process occurs. Specifically, customers have an expectation just prior to the service contact (WE_{ijt-1}), experience a new service contact (DS^*_{ijt}), and develop a posterior prediction of future service (WE_{ijt}). Because customers are *integrating* information, this process implies that both prior information and new information will be positively related to the updated prediction. This logic leads to our first two hypotheses.[3]

H_1: $\partial f_1/\partial WE_{ijt-1} > 0$.
H_2: $\partial f_1/\partial DS^*_{ijt} > 0$.

We believe *should* expectations are influenced from three sources. Similar to *will* expectations, the customer's new *should* expectation (SE_{ijt}) will be related to the customer's prior *should* expectation (SE_{ijt-1}). Second, the *should* expectation may differ between time t and $t - 1$ because of new information reaching the customer between service contacts, such as changes in price, firm communications, and competitive service delivery. We denote this new information as $\mathbf{Z}_{it}$. Third, experiences with the firm's own delivery system can lead to increases, but never decreases, in the customer's *should* expectations between time t and $t - 1$.

An example of the influence of new information is when a firm raises its price and the customers shift their *should* expectations upward to reflect their belief that the service should be better than it was before the price increase. Similarly, if a firm announces that it plans to increase service over previous levels, customers may believe the firm should deliver on this promise. Also, if customers are exposed to a firm's competitor who delivers unanticipated higher levels of service, the customers may believe the firm should deliver similarly high levels of service. For example, Lexus's recent policy of replacing the car when a consumer expresses displeasure with the paint job might alter the consumers' *should* expectations level for other car manufacturers.

We believe that the delivered service influences *should* expectations only when the firm's own service delivery exceeds the individual's prior *should* expectations. Specifically, we postulate that the more the firm's actual delivered service exceeds the customer's prior *should* expectations, the more the customer will increase his or her future *should* expectations for that firm. Thus, in our Lexus example, we would postulate that if the policy of replacing the car exceeds the customer's prior *should* expectations, the customer's *should* expectations for Lexus will increase.

We state these beliefs more formally with the following functional relationship.

$$SE_{ijt} = f_2(SE_{ijt-1}, \mathbf{Z}_{it}, K_{ijt} \cdot DS^*_{ijt}), \quad (2)$$

where $K_{ijt} = 1$ when $DS^*_{ijt} > SE_{ijt-1}$, 0 otherwise. As before, we do not model the $\mathbf{Z}$ vector in any more depth because we control for, but do not measure, these factors.

More specifically, we expect SE_{ijt} to relate directly to SE_{ijt-1}, modified by $K_{ijt} \cdot DS^*_{ijt}$. This leads to our next two hypotheses.

H_3: $\partial f_2/\partial SE_{ijt-1} > 0$.
H_4: $\partial f_2/\partial K_{ijt} \cdot DS^*_{ijt} > 0$.

Equations 1 and 2 make explicit that the two types of expectations are different (albeit related) constructs, and that it should be possible to manipulate one or the other of these expectations via the $\mathbf{X}$ and $\mathbf{Z}$ vectors and different service encounters. We say more on this point in discussing our laboratory study.

Finally, we do not explicitly specify a process that generates *ideal* expectations for two reasons. First, as previously noted, *ideal* expectations are generally unchanged over time; therefore, the *ideal* expectation at time t equals the *ideal* expectation at time $t - 1$. Second, we conjecture that *ideal* expectations influence *should* expectations. The $\mathbf{Z}$ vector in equation 2 could easily include information about an individual's *ideal* expectation.

The Process That Generates Perceptions

We next explicate our conceptualization of how customers form perceptions of the service quality of a firm.[4] This formulation differs from the disconfirmation formulation most often found in the CS/D literature (Oliver 1980) and the gap formulation found in the service quality literature (Parasuraman, Zeithaml, and Berry 1985). However, we show that the implications from our service quality model are compatible with the transaction-specific definition of satisfaction found in the CS/D literature. In addition, we test the viability of our model in relation to the gaps model. In these ways, our model begins to integrate the service quality and satisfaction literatures.

In our model, a person's perception of each of the J dimensions of service quality is conceptualized as a *cumulative* construct, denoted by PS_{ijt}, that is updated each time the person is exposed to the service. We postulate that these perceptions are influenced by a person's ex-

[3]All stated hypotheses are based on the assumption of "all else equal."

[4]Keep in mind that these perceptions are *not* the perceptions of a *specific* service encounter, but instead the perceptions of the service quality based on the consumer's cumulative experience with the firm's service delivery system.

pectations of the service as well as the most recent service encounter. We thus explicitly allow for a person to have a perceptual bias, as our model implies that two customers experiencing an identical service encounter will have different cumulative perceptions of the service if they enter the encounter with different expectations.

Stated more formally, individual i's cumulative perception of the j^{th} dimension of service quality held at time t will be a blend of three factors: the person's expectations just prior to the encounter of what will happen and what should happen, and the new service encounter. The general functional relationship is

$$PS_{ijt} = f_3(WE_{ijt-1}, \mathbf{X}_{it}, SE_{ijt-1}, \mathbf{Z}_{it}, DS^*_{ijt}), \tag{3}$$

where $\mathbf{X}_{it}$ and $\mathbf{Z}_{it}$ are vectors that capture adjustments to expectations occurring between service encounters, as defined in equations 1 and 2.

We believe a person's expectations color the way he or she perceives reality. Specifically, we postulate that customers with higher expectations of what the firm *will* deliver have higher perceptions of the service after an encounter, all else equal, than those with lower *will* expectations. Conversely, customers with higher expectations of what a firm *should* deliver have lower perceptions of the service after an encounter, all else equal, than those with lower *should* expectations. Finally, we believe the delivered service positively affects perceptions. These statements give rise to the following testable hypotheses.

H_5: $\partial f_3/\partial WE_{ijt-1} > 0$.
H_6: $\partial f_3/\partial SE_{ijt-1} < 0$.
H_7: $\partial f_3/\partial DS^*_{ijt} > 0$.

H_5 and H_7 are based on similar logic. We believe customers average/integrate past experience with the firm (which is summarized by their prior *will* expectations) and their latest service encounters in making a cumulative assessment of the service quality level of the firm. This notion leads to our hypothesizing the positive influences. We note that the role of *will* expectations is very similar to the role of the "initial impression" in averaging models of attitude. In these attitude models, initial impressions always have a positive (assimilative) influence.

As distinguished from the assimilative role of the *will* expectation, the *should* expectation acts as a standard of comparison in relation to competitors. As the standard set by competitors goes up, all else equal, the firm fares less well in how it is perceived by customers. Placing our model within the context of assimilation-contrast attitude theory, we are stating that the *should* expectation provides a negative (contrast) influence on overall attitude (perceptions of quality).

Dimensions of Service Quality

A central construct in our model is the customer's perception of overall service quality for a firm. Recent research suggests that this quality assessment is not unidimensional, but instead comprises multiple abstract dimensions (Garvin 1987; Hjorth-Anderson 1984; Holbrook and Corfman 1985; Maynes 1976; Parasuraman, Zeithaml, and Berry 1985; Zeithaml 1988). After studying four consumer service industries, Parasuraman, Zeithaml, and Berry (1985, 1988) identified five dimensions: reliability, assurance, responsiveness, empathy, and tangibles.

We make the assumption that customers perceive the service quality of a system in terms of these five dimensions, and also that their expectations of what *will* and *should* happen are in terms of these five dimensions. We incorporate the multidimensional aspect of overall service quality by defining the following relationship:

$$OSQ_{it} = f_4(PS_{ijt}), \tag{4}$$

where OSQ_{it} equals individual i's overall perception of the firm's service quality at time t, and the j subscript on PS corresponds to the j^{th} dimension of service enumerated by Parasuraman and his coauthors. Note that we postulate that the perceptions of the J different dimensions of the service, and not the "actual" service, directly affect the person's assessment of the overall service quality of a firm. In this way we again acknowledge that perceptual biases are present and that perceptions of reality, not "reality" itself, affect overall attitudes and subsequent behavior.

Previous empirical work suggests that these dimensions of service all have a positive, albeit perhaps unequal, impact on overall quality perceptions. In a variety of different service businesses and industries, respondents consistently rated the dimension of reliability as most important (Parasuraman, Berry, and Zeithaml 1990; Zeithaml, Berry, and Parasuraman 1991). Consistent with previous findings, we believe that though quality is multidimensional, reliability is the key dimension in determining overall perceptions of service quality.

Hence, we hypothesize that the different dimensions of quality are averaged together in some fashion to produce an overall assessment of quality. Further, by substituting equation 3 into equation 4, we can propose hypotheses about the role of the two different expectations and delivered service in customers' judgments of overall quality. Specifically, because the expected signs on PS_{ijt} in equation 4 are positive, we should observe the same direction of effects for the expectation and delivered service constructs as in equation 3. Consequently, we propose the following testable hypotheses.

H_8: $\partial f_4/\partial WE_{ijt-1} > 0$.
H_9: $\partial f_4/\partial SE_{ijt-1} < 0$.
H_{10}: $\partial f_4/\partial DS^*_{ijt} > 0$.

The Relationship Between Overall Quality and Behavioral Intentions

Delivery of high service quality is presumed to relate positively to the success of the firm. Interestingly, no empirical research outside a laboratory setting has been reported that supports this relationship between service quality perceptions and behavioral outcomes of importance to the firm.[5] Unless this positive relationship exists, understanding how customers form judgments about service quality has limited managerial relevance.

We propose the following function to capture this relationship.

$$(5) \qquad BI_{imt} = f_5(OSQ_{it}),$$

where BI_{imt} equals the m^{th} behavioral intention (i.e., loyalty, word of mouth, etc.) for individual i at time t. We strongly believe that service quality positively affects important behavioral outcomes such as loyalty and positive word of mouth. Furthermore, we can substitute through from equations 3 and 4 to examine the indirect effects of expectations and delivered service on behavioral intentions. Because the predicted effect of overall quality in equation 5 is positive, we should observe the same predicted effects of expectations and delivered service as given for equations 3 and 4.

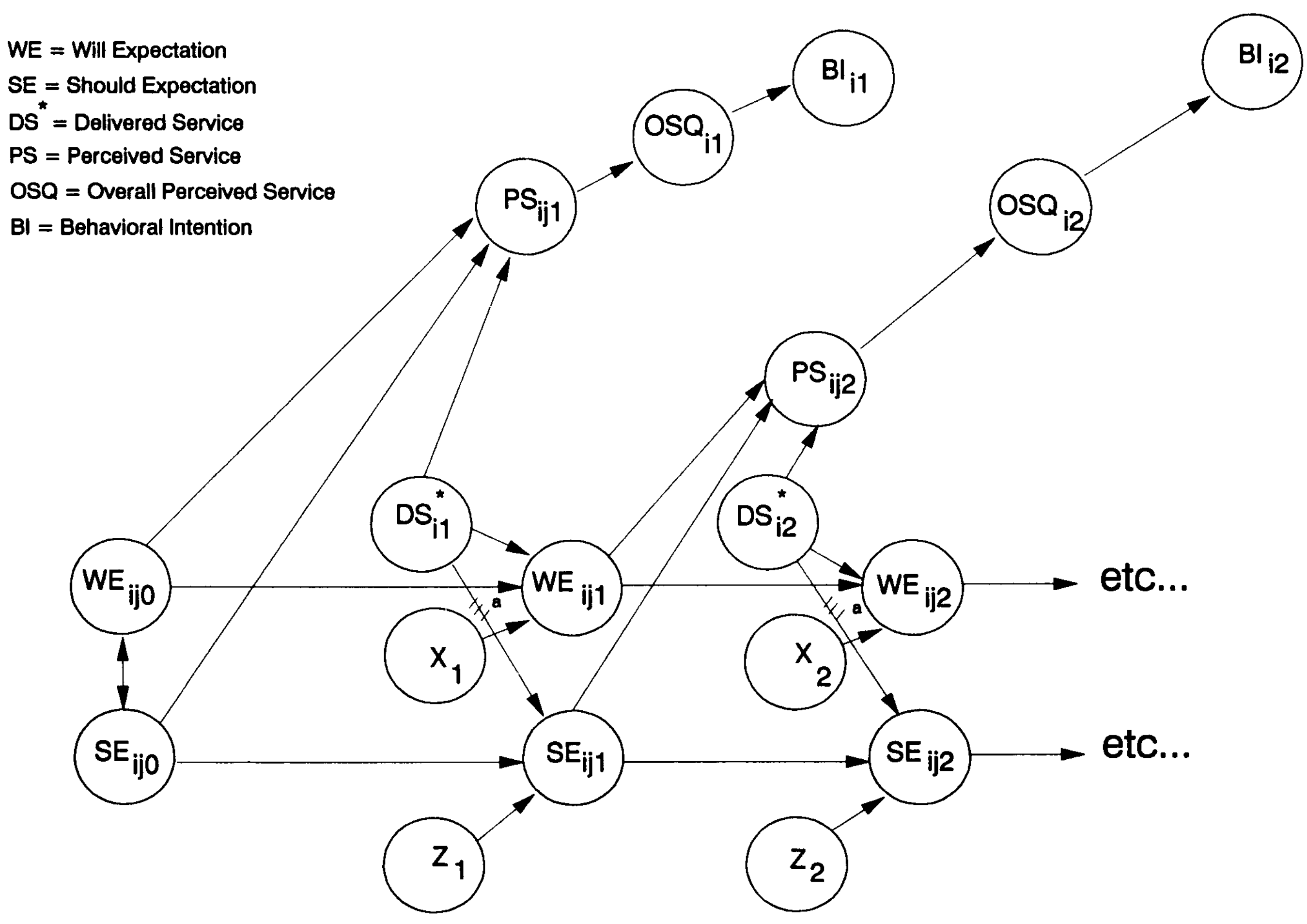
H_{11}: $\partial f_5/\partial WE_{ijt-1} > 0$.
H_{12}: $\partial f_5/\partial SE_{ijt-1} < 0$.
H_{13}: $\partial f_5/\partial DS^*_{ijt} > 0$.

Summary

We present our full conceptual model in Figure 1, which summarizes the proposed relationships among the types of expectations, service quality perceptions, overall perceived service quality, and behavioral intentions. Individuals enter into each service transaction with an initial set of expectations about what *will* and *should* occur on

[5] In the area of customer satisfaction, a recent individual-level study found a significant and positive effect of satisfaction on customer retention (Anderson and Sullivan 1990).

Figure 1
A DYNAMIC PROCESS MODEL OF SERVICE QUALITY



[a] This relationship holds only if $DS^*_{ijt} > SE_{ijt-1}$.

each of the dimensions of service. These initial expectations and the actual delivered service then lead to cumulative perceptions of the delivered service on each dimension, as well as updated expectations for each dimension of what *will* and *should* occur in future transactions. Finally, perceptions of the dimensions of service contribute to an overall assessment of the level of service quality, which in turn leads to behavioral outcomes.

EMPIRICAL TESTING

We now turn to empirical testing of our conceptual model. We begin with an experimental study in which we manipulate the constructs delivered service, *will* expectations, and *should* expectations. We analyze the data in two stages. First, using standard experimental analysis, we use the data to directly test the hypotheses relating to these constructs that emerge from the five functional relationships specified previously. Second, given favorable outcomes of these tests, we specify explicit functional forms for equations 1 through 5. We then estimate this system of equations in a way that enables us to directly test the underlying process suggested by Figure 1. Importantly, this procedure also enables us to directly *test the specification of our explicit functional forms*.

As this specification test provides support for our structural model specification, we next take advantage of this information to develop a method for estimating our dynamic model with data taken at a single point in time. We then apply this approach to data collected in a second study. This second study serves three purposes. First, it increases confidence in the generalizability of our findings because a different data collection approach (a natural field setting), a different service, and a different analytic approach are used. Second, it enables us to test certain relationships that are not explored in study 1. Third, it provides an important insight into how managers can easily and effectively collect data to assess service quality.

STUDY 1

Sample and Data

To test the conceptual model in Figure 1, we first used data obtained from a laboratory experiment involving two simulated visits to a hotel. Subjects were 107 business professionals, including managers and staff, located in a major metropolitan area. Eleven subjects' questionnaires were unusable because of missing data. These subjects were dropped, resulting in a sample size of 96.

Procedure

Subjects participated individually in the experiment and were assigned randomly to one of eight conditions. They were told that the purpose of the task was to find out how they evaluate hotels. A self-administered computer diskette was provided, along with instructions to start the program. Subjects were asked to assume they were to stay at a hotel during a business trip. They were then provided an overview of the task and an explanation of the stimuli. After evaluating a hypothetical restaurant to familiarize themselves with the keyboard and the task, they began the actual task.

Both *should* and *will* expectations were manipulated by providing subjects information about (1) others' perceptions of the quality of target hotel "Alpha" they were to visit and (2) the level of service provided and the price offered by a competitor hotel. Subjects were asked to indicate the quality of hotel service they expected to receive and the quality they thought they should receive. They then "visited" the hotel. After reading a brief general description of their stay, subjects viewed information about the specific performance rating of six features provided by the hotel. Measures of quality assessment were obtained at this point, followed by two behavioral intention measures, and then measures of current *will* and *should* expectations of the level of service.

Subjects were informed they were to stay at the same hotel a second time and were provided information about this second service encounter. Measures of overall quality assessment and the two behavioral intentions and expectations measures were obtained again after this "visit."

Design

We used a three-factor between-subjects design. The factors were initial *will* expectations (three levels: low, medium, and high), *should* expectations (two levels: medium and high), and delivered service (two levels: low and high). Because prior research (Kalra 1992) had indicated difficulty in manipulating *will* expectations to exceed *should* expectations, we focused our attention on obtaining data for only four of the six possible *will-should* pairs (medium *will*/high *should*; medium *will*/medium *should*; high *will*/high *should*; and low *will*/medium *should*). These four conditions were fully crossed with the two levels of service, yielding eight cells. Though incomplete, this design provides unconfounded contrasts for testing of our stated hypotheses (e.g., low *will*/medium *should* vs. medium *will*/medium *should* provides an unconfounded contrast for testing effects of prior *will* expectations). Also, though subjects paid two "visits" to a hotel, we do not use the information relating to the second "visit" in our initial experimental analysis because the first visit affects customers' expectations, thereby negating our ability to conduct planned contrasts for the data related to the second visit. However, we do use this information to test further qualitative implications of the model when we estimate the structural system of equations.

Manipulations

Should and will expectations. On the basis of theory and results of a prior experiment (Kalra 1992), we determined that we could manipulate *should* and *will* expectations by providing information on (1) prior exposure to a competitor's service (which we provided via information on price and rated service of a competitor) and (2) price, word of mouth, and expert opinion on the object hotel (Hotel Alpha). This manipulation was done via a paragraph describing information on the subject's prior visit to a nearby competitive hotel and others' views on Hotel Alpha. Each description held fixed the Hotel Alpha price. As is seen in the discussion on the manipulation checks, this approach resulted in the four desired *will-should* combinations. For more details on these manipulations, see Boulding et al. (1992).

Delivered service. Half of the subjects first received comparatively high performance ratings on the six features provided in their hotel visits, and then comparatively low performance ratings in the second visit. The order was reversed for the other half of the subjects. As we show subsequently, high performance is approximately equivalent to an 80 on our measurement scale and low performance is approximately equivalent to a 50.

Stimuli

Hotels were used as the service setting for two reasons. First, subjects were familiar with the product category. Second, hotels are typically characterized by variability in the quality of service provided during different encounters. The stimuli consisted of a constant neutral description of the stay and manipulated information about performance ratings on six features of the hotel. The features, selected on the basis of a pilot study, were noise, checkin/checkout, amenities, hotel staff, cleanliness, and bed comfort. The subjects were asked to assume that the ratings reflected their own opinions of the features associated with the hotel. Subjects could only view the ratings of the features one at a time. They were free to examine the stimuli as long as they wanted and in any sequence. The performance ratings were displayed in the form of a bar graph anchored between "poor" and "excellent."

Measures[6]

Quality assessments. Quality assessments were measured on a 100-point scale. Subjects were asked to "describe your opinion on the overall quality of Hotel Alpha," which was anchored by the labels "very unfavorable" and "very favorable." The quality assessments were obtained after both the first and second visits. Further, the question was framed so as to obtain measures of overall quality rather than satisfaction/dissatisfaction with a specific visit.

Behavioral intentions. Two questions were asked after each visit about the subject's willingness to provide favorable word of mouth and repeat business. These questions were "How likely are you to stay again at Hotel Alpha?" and "How likely are you to recommend Hotel Alpha to a friend?" Both questions were anchored between "very unlikely" and "very likely." For purposes of analysis, we combine these two questions to form a behavioral intentions scale (equally weighted). This two-item scale yields a Cronbach alpha value of .92.

Expectations. Will expectations associated with the overall service were measured by the question "What is your opinion on the level of service Hotel Alpha will actually provide you?" *Should* expectations were measured by the question "What is your opinion on the level of service you would consider to be reasonable, or Hotel Alpha should provide?" Both questions were framed in terms of future expectations, anchored on one end by "poor service" and on the other "excellent service," and measured on a 100-point scale.

The same order of questions was used in all conditions, which leads to the following sequence of measures for each customer: at $t = 0$, the customer's *will* and *should* expectations; at $t = 1$ and $t = 2$, the delivered service (i.e., high or low), the customer's perceptions of the quality of the service at the hotel, his or her behavioral intentions, and updated *will* and *should* expectations.

Empirical Analysis: Stage 1

Expectation manipulation checks. Table 1 presents cell means for our different measures. The cell means from $t = 0$ for the *will* and *should* expectations enable us to test the effectiveness of our expectation manipulations. Because the delivered service manipulation occurs after the measures taken at $t = 0$, we can collapse the data across the two levels of delivered service. Thus, in the following discussion we combine cells 1 and 5, 2 and 6, 3 and 7, and 4 and 8.

As evidence of the success of our manipulations, the *will* expectation means are higher in cells 2 and 6 than in cells 4 and 8 ($t_{92} = 7.56$) and higher in cells 3 and 7 than in cells 1 and 5 ($t_{92} = 10.37$). As evidence of discriminant validity, the *should* expectations for cells 2 and 6 do not differ from those for cells 4 and 8 ($t_{92} = .65$), nor do cells 1 and 5 differ from cells 3 and 7 ($t_{92} = .00$). Similarly, for the *should* expectation we note the means are higher in cells 1 and 5 than in cells 2 and 6 ($t_{92} = 6.61$). As further evidence of discriminant validity, the *will* expectations for cells 1 and 5 do not differ from those for cells 2 and 6 ($t_{92} = .55$). Consequently, we conclude that (1) we were successful in manipulating initial *will* and *should* expectations into the four desired cells, (2) we were able to identify antecedents to each expectation, and (3) the two different expectations are in fact separate constructs.

Results. Table 1 also gives cell means at $t = 1$ for the two expectations measures, as well as for the measure of overall quality and the behavioral intentions index. A series of cell contrasts enable us to test the qualitative hypotheses generated from the five specified

[6]The measures in this first experiment are at the overall service level instead of the j^{th} dimension level. This is mathematically equivalent to assuming $J = 1$. Given the "limited" exposure to the service, we found subjects were more comfortable with this more macro level of service quality.

Table 1
STUDY 1 CELL MEANS[a]

Manipulations: Expectations *will/should*	Manipulations: Service delivered $t = 1$	Cell	Means of *will expectations* $t = 0$	Means of *will expectations* $t = 1$	Means of *should expectations* $t = 0$	Means of *should expectations* $t = 1$	Means of perceived service quality $t = 1$	Means of behavioral intention index $t = 1$
Med/high	High	1	50.50	64.30	89.70	91.20	62.70	61.55
			(3.69)	(8.85)	(8.72)	(7.42)	(7.99)	(15.71)
Med/med	High	2	51.36	70.00	61.36	77.27	72.27	70.68
			(10.97)	(11.18)	(11.42)	(12.91)	(10.33)	(15.12)
High/high	High	3	84.00	77.42	84.58	84.58	76.16	74.45
			(7.15)	(7.26)	(10.54)	(10.10)	(8.59)	(8.95)
Low/med	High	4	28.75	63.44	61.87	68.43	65.31	63.75
			(18.75)	(9.07)	(20.48)	(18.41)	(12.58)	(10.91)
Med/high	Low	5	54.61	53.46	86.54	83.46	52.69	45.15
			(12.65)	(16.88)	(5.56)	(12.14)	(17.27)	(18.05)
Med/med	Low	6	50.83	59.16	63.75	63.75	55.00	55.21
			(4.88)	(10.83)	(14.94)	(14.32)	(12.43)	(20.35)
High/high	Low	7	87.08	65.41	91.25	86.25	66.67	61.87
			(7.52)	(10.54)	(8.01)	(13.16)	(9.37)	(10.45)
Low/med	Low	8	26.00	39.00	57.50	66.00	48.00	37.50
			(9.07)	(11.50)	(14.95)	(14.10)	(11.83)	(14.67)

[a]Standard deviations are in parentheses.

functional relationships.[7] Table 2 indicates the relevant cell contrasts that hold "all else equal" for tests of our hypotheses, as well as the test results. As a check on our specified equations, this table also includes tests of significance for relationships not included in the equations.

Because the specified functional relationships do not contain interaction terms (with the exception of the *should* expectation equation), we first conduct joint tests of significance for all interaction terms estimable from our experimental design for equations 1, 3, and 5. We find no joint interaction effects in any of these three equations reaching significance at the .10 level.[8] Therefore, for these equations, we restrict our attention to testing for main effects.

Starting with the *will* expectation updating equation, Table 2 indicates strong support for our two main effects hypotheses. Results of the cell contrasts suggest that prior *will* expectations and the delivered service both contribute positively and significantly to the updated *will* expectations, supporting H_1 and H_2. We also find marginal significance ($p = .10$) for the effect of prior *should* expectations in the *will* expectation equation. Because we do not hypothesize this relationship, we explore it in greater depth in the stage 2 analysis of our experimental data.

Turning to the *should* expectation updating equation, we note that the prior *should* expectation significantly and positively affects the updated *should* expectation (H_3). Consistent with our specification of equation 2, we find no main effect of delivered service, nor do we find an effect of prior *will* expectations. In addition, looking at differences in values between $t = 0$ and $t = 1$, we find significant support for H_4, that is, delivered service increases *should* expectations, but only if the delivered service exceeds the prior *should* expectation. This occurred only in cells 2 and 4, where the delivered service was high (approximately 80) and the prior *should* expectations were medium (approximately 60).

Table 2 also indicates strong support for the three main effects hypotheses relating to our perceived service equation (assuming a single dimension of quality). In support of H_5, prior *will* expectations positively and significantly influence cumulative assessments of perceived service. In support of H_6, prior *should* expectations negatively and significantly affect perceived service. Also, as predicted by H_7, delivered service positively and significantly contributes to perceptions of service quality.

Finally, it is interesting to see whether our manipulations affect the downstream behavioral intentions measures. Table 2 again provides support for all of our stated main effects hypotheses. As suggested by H_{11} and H_{12}, prior *will* expectations positively and significantly influence behavioral intentions, whereas prior *should* expectations negatively and significantly influence behavioral intentions. As one would expect, and as stated by H_{13},

[7]As noted in footnote 6, in the context of this experiment we assume a single dimension of quality. We therefore cannot distinguish between the constructs "perceptions" (*PS*) and "overall quality" (*OSQ*) as laid out in equations 3 and 4. Thus, equation 3 and H_5–H_7 are redundant with equation 4 and H_8–H_{10}. In study 2 we generate separate estimates of equations 3 and 4.

[8]In the *will* expectation equation, $F_{3,88} = 2.14$; in the perceived service equation, $F_{3,88} = .80$; and in the behavioral intentions equation, $F_{3,88} = 1.02$.

Table 2
HYPOTHESIS TEST RESULTS[a]

Hypothesis[b]	Dependent measure	Factor	Cell contrast	t-statistic[c]
H_1	WE_t	WE_{t-1}	1,5 vs. 3,7	4.08*
H_1	WE_t	WE_{t-1}	2,6 vs. 4,8	3.24*
H_2	WE_t	DS_t^*	1,2,3,4, vs. 5,6,7,8	5.99*
NH	WE_t	SE_{t-1}	1,5 vs. 2,6	1.88***
H_3	SE_t	SE_{t-1}	1,5 vs. 2,6	4.17*
NH	SE_t	WE_{t-1}	1,5 vs. 3,7	.36
NH	SE_t	WE_{t-1}	2,6 vs. 4,8	.70
NH	SE_t	DS_t^*	1,2,3,4 vs. 5,6,7,8[d]	1.30
H_4	$SE_t - SE_{t-1}$	$K_t \cdot DS_t^*$	2,4 vs. 1,3,5,6,7,8	3.39*
H_5	PS_t	WE_{t-1}	1,5 vs. 3,7	4.16*
H_5	PS_t	WE_{t-1}	2,6 vs. 4,8	1.36***
H_6	PS_t	SE_{t-1}	1,5 vs. 2,6	1.78**
H_7	PS_t	DS_t^*	1,2,3,4 vs. 5,6,7,8	5.44*
H_{11}	BI_t	WE_{t-1}	1,5 vs. 3,7	4.26*
H_{11}	BI_t	WE_{t-1}	2,6 vs. 4,8	2.14**
H_{12}	BI_t	SE_{t-1}	1,5 vs. 2,6	2.92*
H_{13}	BI_t	DS_t^*	1,2,3,4 vs. 5,6,7,8	6.11*

[a]Given directional hypotheses, significance tests are one-tailed. For nonhypothesized relationships, significance tests are two-tailed.
[b]NH indicates not hypothesized.
[c]Calculated t-statistics are based on 88 degrees of freedom.
[d]This is a very conservative test. A less conservative test would contrast cells 1,3 vs. 5,7.
*Significant at the .01 level.
**Significant at the .05 level.
***Significant at the .10 level.

delivered service positively and significantly influences behavioral intentions.

In summary, our experimental data provide strong support for all of our hypothesized relationships given by equations 1 through 5 and all but one of the implicit null hypotheses. We highlight three conclusions from these analyses. First, we demonstrate conclusively that *will* and *should* expectations are different constructs. Second, we confirm that *will* expectations positively influence and *should* expectations negatively influence perceptions of quality. Third, holding fixed the delivered service, one can trace a measurable effect of these two different expectations on individuals' behavioral intentions.

The preceding univariate contrasts take full advantage of the experimental design aspect of our data. Further, such contrasts require no assumptions about the specific functional forms of equations 1 through 5. A drawback of this approach, however, is that it does not allow us to estimate the magnitude of the effects. It also does not allow us to specify and test any cross-equation implications of our process model as given in Figure 1, or use the period 2 data for estimation. To accomplish these latter objectives, we must make specific assumptions about the functional forms of equations 1 through 5. However, we can directly test those assumptions. As long as these assumed functional forms are costless (i.e., they are not rejected by specification tests), we favor this approach because it yields the benefit of additional specificity.

Empirical Analysis: Stage 2

Specification of functional forms. In specifying our process model, we focus our attention on the functional relationships for our measured constructs *WE*, *SE*, *PS*, and *BI*. We do not specify the role of "other" information (i.e., the **X** and **Z** vectors) in this process, because in our studies we control for (but never measure) "other" information.[9] Consequently, we leave this part of our conceptualization to future research.

With the preceding caveat in mind, we now fully specify equation 1 (we distinguish the fully specified equation by "S":

$$(1S) \qquad WE_{ijt} = \alpha_{jt} WE_{ijt-1} + (1 - \alpha_{jt}) DS_{ijt}^* + \epsilon_{1it},$$

where the parameter α_{jt} (which can assume values between 0 and 1) determines the relative weights assigned to the prior expectation and the delivered service in the updating equation, and ϵ_{1it} is an error term. This specification makes equation 1S consistent with our conceptualization of *will* expectations being updated by means of an averaging process.

[9]In study 1, no information is provided between our measurement of WE_{t-1} and the service encounter at time t. In study 2, we explicitly control for differences in the informational environment by conducting within-person analysis.

We note two testable implications of this specification. First, one can test the validity of the averaging model by comparing the estimates of equation 1S, which contains a single constrained parameter, with an additive model, which would contain two unconstrained parameters. Second, there is a qualitative implication as to the value of α_{jt} over time. Specifically, as an individual accumulates more information about a service on the basis of past experience, the influence of WE_{ijt-1}, which captures this past experience, should receive more weight in relation to the most recent service encounter in generating future predictions about the service. Consequently, α_{jt} should grow larger over time.

Turning to the *should* expectation updating process, we fully specify equation 2 as

$$(2S) \qquad SE_{ijt} = SE_{ijt-1} + \beta_{jt}(K_{ijt} \cdot DS^*_{ijt}) + \epsilon_{2it},$$

where β_{jt} equals an updating parameter postulated to be greater than zero and ϵ_{2it} is an error term. Unlike equation 1S, this updating formulation does *not* follow an averaging process. Instead, we believe *should* expectations follow a ratcheting process—they can go up, but they cannot go below prior levels.[10] More specifically, equation 2S implies that *should* expectations equal old *should* expectations unless the latest delivered service exceeds the prior *should* expectation.

For the perception equation, we make the assumption that customers blend their prior *will* expectations with the latest service experience in a manner identical to that used to form the new *will* expectation. This averaging process is then additively blended with the consumer's *should* expectations. We represent these beliefs mathematically as

$$(3S) \quad PS_{ijt} = \alpha_{jt}WE_{ijt-1} + (1 - \alpha_{jt})DS^*_{ijt} + \gamma_{jt}SE_{ijt-1} + \epsilon_{3it},$$

where α_{jt} is the same updating parameter as given in equation 1S, γ_{jt} is a new updating parameter with $\gamma_{jt} < 0$, and ϵ_{3it} is an error term. Such a formulation makes explicit the assimilative (averaging) role of prior *will* expectations and the latest delivered service, and the contrastive (adding) role of *should* expectations.[11]

As with equation 1S, we can test the explicit averaging assumption of 3S by comparing the constrained one-parameter averaging process with an unconstrained two-parameter process. We can also test whether the parameter α_{jt} is the same parameter as given in equation 1S by comparing an unconstrained estimate of this parameter with one that is constrained to equality across equations 1S and 3S.

Note that by replacing the first two terms of equation 3S with the left side of equation 1S, one can see the difference between an individual's current *perception* of the service and his or her current *prediction* of likely future service. Specifically, these two constructs differ increasingly as the consumer's prior expectations of what service they *should* get increase. Two implications are that we expect the parameter on the *should* expectation term to be significant in equation 3S, but not to add explanatory power in equation 1S. We test both implications with all of our longitudinal data.

Because our conceptualization of the role of expectations in generating perceptions is new, we use simple examples to reconcile it with previous conceptualizations and demonstrate its intuitive appeal. First, for notational simplicity, assume that there is only one dimension of service quality. Further assume that two individuals have identical expectations of what should happen but, on the basis of prior service, the first individual predicts future service will rate a five and the second individual predicts future service will rate a seven. Given our system of equations, the implication is that individual 2's current perception is higher than individual 1's. Finally, suppose these two individuals simultaneously experience a new service encounter and receive identical delivered service that rates a six. We next explore three questions:

- Which of these two individuals will have higher cumulative perceptions of the quality of the firm's delivery system after the encounter?
- Which will be most satisfied with the most recent encounter?
- How are the perceptions of each individual changing over time?

According to the hypothesized relationships in equation 3S, the second individual will have a higher perception of the service because this individual combines all prior experience with the latest transaction to form a cumulative perception of the service. However, on the basis of a definition of satisfaction associated with the disconfirmation of expectations paradigm (i.e., satisfaction equals performance minus expected performance, or $DS^*_{ijt} - WE_{ijt-1}$ in our notation), the first person will be more satisfied with the most recent service encounter because this person had the lower prior *will* expectation. Finally, given the form of equation 3S, the second individual (who originally had the higher value for overall perceived quality) will show a decrease in his or her cumulative perception of quality because the last service encounter was below his or her predicted level of service. Similarly, the first individual (who had the lower value for overall perceived quality) will show an increase in cumulative perceived quality because his or her last service encounter exceeded his or her predicted level of service. In this way, our model is compatible with the CS/D disconfirmation of expectations paradigm and

[10]Conceptually, we believe that information in the **Z** vector could cause these expectations to decrease. For example, a price decrease by the firm or learning about much higher levels of price for competitors could lower *should* expectations.

[11]An alternative theoretical basis for 3S is that *will* expectations affect the private psychological impression of the service, whereas the *should* expectations affect the output mapping of private psychological impressions onto overt ratings of perceptions (Lynch, Chakravarti, and Mitra 1991).

shows how satisfaction with a specific transaction can lead to an increase in perception of the overall service.

In a similar vein, assume two individuals with identical *will* expectations, but different *should* expectations, receive identical actual service. In this case, our model predicts that the individual with the higher expectation of what the service *should* deliver perceives the service more negatively. One way of thinking of this effect is that some individuals are more critical or demanding than others. Individuals who develop higher expectations of what service *should* occur (perhaps because of firm or competitive actions) are more difficult to please (i.e., they form more negative evaluations) than individuals with lower *should* expectations.

Finally, our model has the pleasing characteristic that improvements in the service delivery system lead to more positive service quality perceptions.[12] To see this effect, imagine a firm engaged in "continuous quality improvement," a currently popular business concept. If the *will* expectation itself had a negative effect, and this expectation increased with the increasing service, a firm would have to constantly upgrade its delivered service simply to stay even with the rising expectations. Said differently, despite improved service delivery, there would be no resulting improvements in consumer perceptions of quality. In contrast, our model suggests that firms benefit from increasing consumer expectations of what service will occur. Specifically, if a firm increases expectations of what will occur by improving delivered service, the consumer will perceive a higher quality service than before the service improvements occurred, even if the service plateaus at this higher level.[13]

We next specify a simple functional form suggesting that overall perceived quality is a linear combination of perceptions of different dimensions of quality.

$$OSQ_{it} = \sum_{j=1}^{5} \phi_j PS_{ijt} + \epsilon_{4it}, \quad (4S)$$

where the ϕ_j are > 0 and ϵ_{4it} represents an error term.

Our dynamic formulation as stated by equations 1S through 4S differs from the prevailing model of service quality, the gaps model (Parasuraman, Zeithaml, and Berry 1985). We label the gaps model a static formulation because no temporal sequencing is specified. In the static model, overall quality is represented by gap 5, which consists of perceptions minus contemporaneous *should* expectations. Gap 5 is in turn a function of gaps 1 through 4, which contain, among other things, information about actual service. Representing the static model relationship between overall perceived quality and gap 5 in equation form results in the following expression.

$$OSQ_{it} = \sum_{j=1}^{5} \Theta_j(PS_{ijt} - SE_{ijt}) + \epsilon_{6it} \text{ (the gaps model)} \quad (6)$$

Two points should be noted about equation 6. First, because the gaps model is static, expectations prior to the service contact are assumed implicitly to equal those after the service. Thus, equation 6 is stated in terms of contemporaneous perceptions and expectations. Second, the gaps model implicitly constrains the parameters on perceptions and expectations to negative equality. A more general formulation of equation 6 allows these coefficients to vary, that is,

$$OSQ_{it} = \sum_{j=1}^{5} (\Theta_{1j} PS_{ijt} + \Theta_{2j} SE_{ijt}) + \epsilon_{7it} \quad (7)$$

(the generalized gaps model).

Note that our dynamic specification, given by equation 4S, is also a special case of equation 7. In particular, our dynamic model formulation implies $\Theta_{1j} > 0$ and $\Theta_{2j} = 0$. In contrast, the gaps model predicts $\Theta_{1j} = -\Theta_{2j}$.

We put forward these alternative specifications of the overall quality equation to highlight our differences with the prevailing model of service quality. Because we do not obtain separate measures of perceptions and overall quality, we cannot test the different specifications by using our study 1 data. However, in study 2 we compare the validity of our formulation represented by equation 4S with the specifications given by equations 6 and 7.

Finally, we specify behavioral intentions as a simple linear function of overall quality perceptions.

$$BI_m = \lambda_m OSQ_{it} + \epsilon_{5it}, \quad (5S)$$

where $\lambda_m > 0$ and ϵ_{5it} is an error term.

Empirical model and results. We have three longitudinal measures of the *will* and *should* expectations and two measures each for the perceived service, the actual service delivered, and the behavioral intentions. Therefore, we can estimate our system of equations (i.e., equations 1S, 2S, 3S, and 5S) at two different time periods. Before doing so, however, we need to relate our obtained measures to the theoretical constructs in our structural model. In particular, we need to calibrate all of our measured constructs onto a common scale. We do this by making the arbitrary, but nonrestrictive, assumption that our measured expectations, denoted by *M* preceding the expectation notation (i.e., *MWE* and *MSE*), establish the metric for our empirical measures. Thus, we scale all of the other measures in relation to the measured expectations variables. We present the details of this scaling calibration in Appendix A.

The scaling calibrations enable us to write our structural model in terms of measured variables. The resultant empirical model includes four scaling parameters—*a*, *b*, *c*, and *d*—in addition to our structural parameters. Substituting equations A1 through A3 from Appendix A into equations 1S, 2S, 3S, and 5S yields the following empirical model.

[12] This is also a characteristic of the conceptual model of Bolton and Drew (1991a).

[13] Technically, if delivered service plateaus, that is, $DS_t^* = DS_{t-1}^*$, our model yields more positive future service quality perceptions if $(1 - \alpha) > -(\gamma\beta K)$.

(8) $MWE_{it} = (1 - \alpha_t)a + \alpha_t MWE_{it-1} + (1 - \alpha_t)bMDS^*_{it} + \epsilon_{8it}$

(9) $MSE_{it} = \beta_t aK_{it} + MSE_{it-1} + \beta_t bK_{it} \cdot MDS^*_{it} + \epsilon_{9it}$

(10) $MPS_{it} = -c + (1 - \alpha_t)a + \alpha_t MWE_{it-1} + (1 - \alpha_t)bMDS^*_{it} + \gamma_t MSE_{it-1} + \epsilon_{10it}$

(11) $MBI_{it} = -d + \lambda_t c + \lambda_t MPS_{it} + \epsilon_{11it}$

where t = 1, 2 for equations 8 through 11, yielding an empirical model composed of eight equations. Two points must be made about these equations. First, the coefficients of this empirical model often confound the structural and scaling parameters. For example, the intercept term in equation 8 is the product of the structural parameter $(1 - \alpha_t)$ and the scaling parameter a. Even so, we are able to obtain at least one direct estimate of each of the structural parameters. Second, unlike the structural model parameters, none of the scaling parameters are time-subscripted because we see no reason to expect our theoretical construct scales to shift over time.

We have three goals in estimating this empirical model. First, we want to test the validity of the complete specification of our process model. Second, we want to confirm our previous qualitative findings by using a second analytic approach and an extended dataset. Finally, we want to extend our insights by examining how the information integration parameters change over time.

We start by addressing the specification issue. We do this by first estimating our empirical model by using OLS, imposing no restrictions across or within equations. Thus, for example, we do not constrain the parameters on MDS^*_{it} in equations 8 and 10 to equal one minus the parameter on MWE_{it-1} (in either equation 8 or equation 10) times the scaling parameter b. Estimation in this manner yields estimates of 22 coefficients. As noted previously, most of the coefficients in equations 8 through 11 do not provide direct estimates of our structural parameters. However, this procedure does yield two unconfounded estimates each for the structural parameters α_1 and α_2 and one direct estimate each for the structural parameters γ_1, γ_2, β_1, β_2, λ_1, and λ_2. These results, along with the confounded estimates, are reported in Table 3.

We make two observations about the results in Table 3. First, we note that all of the unconstrained, direct estimates of our structural parameters have the predicted direction as well as statistical significance. Second, though we subsequently formally test for equality, we note that the two independent unconstrained estimates of α_1 and α_2 (i.e., the coefficients on MWE_{t-1} in equations 8 and 10) are approximately equal. Table 3 indicates that the two estimates of α_1 equal .34 and .38 and the two estimates of α_2 equal .80 and .87.

We now formally test the numerous within- and cross-equation constraints, that is, assumptions implied by our conceptual model. An example of a cross-equation constraint is that the coefficients on MWE_{t-1} in equations 8 and 10 should be equal. An example of a within-equation constraint is forcing the three α parameters in equation 8 to equality. We impose these constraints by using SAS's nonlinear search procedure, SYSNLIN. More technically, we estimate only the four parameters (at two different points in time) of theoretical interest, α_1, α_2, γ_1, γ_2, β_1, β_2, λ_1, and λ_2, and the four scaling parameters a, b, c, and d. This results in reducing a 22-parameter model to a 12-parameter model. Because the constrained model is a subset of the unconstrained model, we can perform a likelihood ratio test based on the sums of squared errors (Amemiya 1985). This ratio is distributed

Table 3

UNRESTRICTED MODEL ESTIMATES OF EQUATIONS 8–11 FOR t = 1 AND t = 2[a]

	Dependent variables							
	Eq. 8		*Eq. 9*		*Eq. 10*		*Eq. 11*	
Independent variables	$MWE_{t=1}$	$MWE_{t=2}$	$MSE_{t=1}$	$MSE_{t=2}$	$MPS_{t=1}$	$MPS_{t=2}$	$BI_{t=1}$	$BI_{t=2}$
Constant	35.79*	.33	.00	.41	47.88*	9.63	−3.41	4.73***
	(3.009)	(6.135)	(1.652)	(.961)	(4.807)	(8.095)	(5.338)	(2.667)
MWE_{t-1}	.34*	.80*			.38*	.87*		
	(.046)	(.086)			(.059)	(.113)		
MDS_t	15.00*	21.13*			14.20*	22.48*		
	(2.170)	(2.559)			(2.211)	(3.000)		
$K^*MDS^*_t$			10.37*	8.69*				
			(3.115)	(2.007)				
MSE_{t-1}			1.00[b]	1.00[b]	−.17**	−.24*		
			(0.0)	(0.0)	(.075)	(.093)		
MPS_t							.99*	.93*
							(.083)	(.046)
R^2	.504	.533	.288	.624	.465	.463	.602	.814

[a]Standard errors are in parentheses.
[b]The dependent measure in this equation equals $(MSE_t - MSE_{t-1})$.
*Significant at the .01 level.
**Significant at the .05 level.
***Significant at the .10 level.

Table 4
RESTRICTED MODEL ESTIMATES OF EQUATIONS 8–11 FOR t = 1 AND t = 2[a]

Parameter	*Estimate*
α_1	.38* (.035)
α_2	.55* (.044)
β_1	.13* (.034)
β_2	.12* (.023)
γ_1	−.11** (.055)
γ_2	−.21* (.055)
λ_1	.91* (.042)
λ_2	.98* (.042)
a	49.18* (1.944)
b	29.41* (2.941)
c	−9.94** (4.226)
d	−11.24** (4.545)

[a]Standard errors in parentheses.
*Significant at the .01 level.
**Significant at the .05 level.

χ^2 with degrees of freedom equal to the number of imposed restrictions. The test enables us to identify whether our structural model imposes binding constraints, that is, whether we have misspecified the process by which our subjects form perceptions of quality, which in turn leads to behavioral intentions.

We find that the cross- and within-equation constraints are *not* binding. The estimated χ^2 equals 2.46, which is well below the critical χ^2 value for p = .5 and 10 degrees of freedom (χ^2_{10} = 9.34).[14] Thus, we come nowhere close to rejecting the underlying assumptions embedded in our process model. Though we recognize that this finding does not mean that we have specified the best possible process model, it does increase our confidence in the robustness of our specification.

Table 4 presents the fully constrained estimates of our structural model. In discussing these estimates, we first note that all of the estimates of the eight structural parameters of theoretical interest are significant and in alignment with our qualitative hypotheses. In addition, because we now obtain estimates of the scaling parameters, we note that on the 100-point expectations scale, "low" service is estimated to equal 49.2 and "high" service is estimated to equal 78.6 (i.e., a = 49.2 and b = 29.4).

For patterns of results over the two periods in the data, we note the increase in α_t from period 1 to 2 (from .38 to .55). This result is compatible with our hypothesis that this parameter will grow over time as customers accrue more prior experience and weight the current experience less in updating their *will* expectations and perceptions of service quality. We also note the doubling in absolute size in γ, the coefficient on the *should* expectation in the perception equation, from period 1 to period 2 (from −.11 to −.21). Though not hypothesized, this result relates to a specification issue we now address.

In particular, as another check on our specification, we include the prior *should* expectation in the *will* expectation updating equation. When we estimate this equation for time = 1, we find a significant coefficient on the *should* expectation variable, as one would expect given our ANOVA findings. However, at time = 2 this coefficient loses significance (t_{92} = 1.07). This result, coupled with the observation that γ_t increased in absolute size from −.11 to −.21 over the two time periods, leads us to conclude that the strength of the customers' beliefs about what should happen became stronger as customers gained more exposure to the service situation. Consequently, their beliefs about perceptions of quality became more distinct from their beliefs about what will happen. In particular, the *should* expectation loses significance in equation 1S (where it is hypothesized to have no effect) and becomes more significant in equation 3S (where it is hypothesized to have a significant negative effect).

Finally, perceptions of quality strongly influence behavioral intentions. Interestingly, the parameter capturing this effect, λ, increases in size from period 1 to period 2 (.91 to .98). Though a marginal increase in size, this result warrants further investigation because it implies that over time perceptions of quality become increasingly important in driving behavioral intentions.

Discussion

The results reported in stage 1 and stage 2 of study 1 are very compatible with our postulated process model. We interpret these results as providing strong evidence that a person's prior *will* and *should* expectations, and the delivered service, influence a person's perceptions of quality. Moreover, our results provide strong support for our conceptualization that *will* expectations positively influence perceptions of quality and *should* expectations negatively influence perceptions of quality. These perceptions, in turn, positively influence behavioral intentions.

Perhaps of greater significance is the finding that the numerous within- (e.g., averaging of *will* expectations and delivered service in the *will* expectation updating equation) and cross-equation (e.g., the same terms appearing in the *will* expectation updating equation and the

[14]Because we want to *accept* the null hypothesis that our structural model does not impose binding constraints, we use a p-value of .5 to increase the power of rejecting the null hypothesis.

perception equation) restrictions were nonbinding. Rather than assuming we correctly specify these equations, or simply accepting the individual coefficient estimates from these equations as compatible with our conceptualization, we directly test the "assumptions" (implied restrictions) inherent in our model. Thus, we take the unusual step of simultaneously testing the significance of the structural parameters and subjecting our structural model to a stringent specification test. *Both* of these test results strongly support our conceptualization.

Given these results, we next address the generalizability of our model. We do this by presenting the results of a second study in which a very dissimilar research method (a field study) was used to assess customers' perceptions and expectations of a real service.[15] Compatible results with this second approach would provide substantial evidence that our obtained results are not due to some confound (unknown to us) associated with the laboratory experiment.

STUDY 2

Besides providing a vehicle for replicating our previous finding by a very different research approach, study 2 provides two additional sources of value. First, it enables us to estimate equations 4S, 6, and 7, which we could not do in study 1. Second, it provides a setting to explicate a research method that enables firms to track customers' evaluations of service quality, an inherently dynamic process as shown by our model, using data collected from a single point in time. This method is exceptionally valuable in that it enables managers to infer the impact of their delivered service *without measuring* this construct. Therefore, these explorations should be of great interest to practioners involved in measuring service quality and cumulative customer satisfaction.

Sample

The data for this study were obtained from a major study of service quality of an educational institution. This study, commissioned by the top management of the institution, was based on 177 obtained responses of the current customers of the institution. Participation by the customers was both voluntary and confidential, resulting in a 46% response rate. Monetary rewards were provided to a small set of participating customers as a result of a lottery drawing.

Operationalization of Variables

Expectations and perceptions of the five dimensions of service quality were measured by 36 statements taken from SERVQUAL (Parasuraman, Zeithaml, and Berry 1988) and then modified by top managers and the research team to capture more precisely expectations and perceptions associated with an educational service. Respondents recorded these expectations and perceptions by indicating their agreement with each statement on a 1 to 7 scale. Approximately half of the sample gave expectations data in the form of what customers thought *should* happen, whereas the other half of the sample gave their predictions of what *will* happen in the education process. Using the original SERVQUAL scales as a guide, and after performing factor analysis in combination with managerial judgment, we grouped the 36 questions to form multiple measures for the five dimensions of service quality (reliability, responsiveness, assurance, empathy, and tangibles). Table 5 provides the scale for two of the five dimensions, along with the Cronbach alpha values for the perception and expectation scales. Boulding et al. (1992) report this information for all five dimensions.

Table 5

EXPECTATION AND PERCEPTION SCALE ITEMS[a]

Reliability
Professors and teaching assistants will grade fairly and accurately.
Courses will be well taught.
The staff will ensure that the MBA program runs smoothly.
Professors will be organized and prepared for class.
When professors promise to be available during office hours, they will be there to see students.
Professors will have prior teaching experience before coming to this organization.
Cronbach alpha:
Expectations = .74
Perceptions = .73
Empathy items
Professors will give students individual attention.
Professors will help students with personal problems and career advice.
Students will be able to contact a professor at home.
Professors will know what the needs of their students are.
Professors will have their students' best interest at heart.
Cronbach alpha:
Expectations = .69
Perceptions = .74

[a]*Should* expectations substituted the word "should" for "will." Perceptions substituted the word "are" for "will." Also, we report a single Cronbach alpha for expectations because the *will* and *should* contemporaneous expectations differ only by an individual-specific constant.

The alpha values indicate that the reliability of all our scales equals or exceeds .60, with most exceeding .7. On the basis of this result and prior evidence of Parasuraman, Zeithaml, and Berry (1988) that there are five dimensions of service quality, we formed indices by averaging the responses to the individual measures associated with a dimension. These indices are used as summary measures of the five underlying constructs in estimating equations 4S, 6, and 7.

Overall quality of the educational service and six items of intended individual-level behaviors of strategic importance to the school were also measured. The latter measures included such items as saying positive things

[15]This "second" study was actually conducted and analyzed prior to the first (laboratory) study. However, the analysis scheme used for the second study relies on many of the assumptions tested in the laboratory study. Hence, for exposition, we reverse the order of discussion for the two studies.

about the school to people outside the school, planning to contribute money to the class pledge upon graduation, and planning to recommend the school to one's employer as a place to recruit. These six behavioral intention variables were grouped into a single index measure. As evidence that these six items tap the underlying behavioral intentions construct, the Cronbach alpha is equal to .80. Finally, country of origin, gender, degree of prior work experience, first or second year in program, and area of educational concentration were measured to partially control for individual differences in the subsequent analyses. As we have no theoretical interest in these variables, we do not discuss them further, though they are included during estimation as covariates.

Estimation

Our data are similar to most service quality field study data in that they consist of a cross-section of self-reported information taken at one point in time. Consequently, they impose some limitations on direct estimation of the structural equations developed in the preceding section. Specifically, because we measure expectations variables contemporaneously with perceptions variables, we do not have measures of prior expectations.[16] Finally, we do not measure actual delivered service because any measure obtained from the customer immediately becomes a perception of the service, and because we were unable to match objective organization measures of actual service to the individuals receiving the service. This is a typical problem in any service setting.

Even though our particular data constrain us from directly estimating equations 1S, 2S, and 3S, we show that it is still possible to obtain consistent estimates of the two key parameters in our process model, that is, α_{jt} and γ_{jt}, as well as the parameters in equations 4S, 5S, 6, and 7, simply by using contemporaneous measures of all the relevant constructs. Moreover, these two parameter estimates enable us to perform falsifying tests on our basic structural model, as well as gain insight into the relative influences of prior *will* and *should* expectations *and* the unobserved delivered service on customers' current perceptions of the cumulative level of service quality of the five dimensions of service quality.

More specifically, we derive consistent estimates for α_{jt} and γ_{jt} from reduced-form equations that assume equations 1S, 2S, and 3S represent the true underlying process. We justify this assumption on the basis of three points. First and foremost, results from our study 1 tests indicate that equations 1S, 2S, and 3S are an excellent representation of the service quality process. Second, using our study 2 data, we can directly test two implications from our model by looking at the signs and magnitudes of the estimates of the two structural parameters α_{jt} and γ_{jt}. Empirical results inconsistent with hypothesized signs and magnitudes for these coefficients would indicate the nonviability of our postulated structural model. Third, we cannot come up with a plausible alternative structural model that could produce the results generated in our reduced-form modeling approach.

Our approach is to derive reduced-form equations that are in terms of only observable contemporaneous variables. We start by specifying one additional nonrestrictive relationship between our two expectation constructs. Note that in a field setting there is normally a strong cross-sectional correlation between SE_{ijt} and WE_{ijt}, even though equations 1S and 2S do not impose any direct contemporaneous relationship between these two constructs. We capture this correlation by specifying the relationship between SE_{ij} and WE_{ij} for any arbitrary point in time, t, to be

$$(12) \qquad SE_{ijt} = WE_{ijt} + \mu_{ijt},$$

where μ_{ijt} captures all individual differences at time t. Thus, μ_{ijt} includes the influence of the individual's previous service experiences, factors included in the **X** and **Z** vectors for the past t periods, and individual differences associated with the degree to which a customer is critical or demanding.

Appendix B presents the derivation of our estimating equations. This derivation consists of three simple steps. First, we use equation 12 to rewrite equation 3S completely in terms of either the *will* or the *should* expectation. We take this step because we measure either the *will* or the *should* expectation for each individual, but not both. Second, we use equations 1S and 2S to write these equations in terms of contemporaneous expectations. Third, by recognizing that we have N measures for each individual on each of the j dimensions of service quality (see Table 5), we utilize the multiple (repeated) measures aspect of our design and thereby remove (control for) all of the factors that are fixed for a specific individual (i.e., delivered service). This is done by "mean-differencing" the data, as shown in Appendix B.

These three steps yield the following two reduced-form equations.

$$(13) \qquad (PS_{ijnt} - PS_{ij\cdot t}) = \left(\frac{\alpha_{jt} + \gamma_{jt}}{\alpha_{jt}}\right)(WE_{ijnt} - WE_{ij\cdot t}) + \epsilon_{13it},$$

$$(14) \qquad (PS_{ijnt} - PS_{ij\cdot t}) = (\alpha_{jt} + \gamma_{jt})(SE_{ijnt} - SE_{ij\cdot t}) + \epsilon_{14it},$$

where PS_{ijnt}, WE_{ijnt}, and SE_{ijnt} are the n^{th} measure of the j^{th} dimension for the appropriate construct and the • notation indicates the mean for the i^{th} individual on the j^{th} dimension.

In words, equations 13 and 14 state that for each individual there is a relationship between how that person responds to the n^{th} perception question tapping dimension j relative to his or her mean response and how the

[16]To test for response biases due to priming with either the *will* or *should* expectation, we obtain measures of either the contemporaneous *will* expectation or the *should* expectation, but never both from the same respondent. As will become evident, this decision to ask for either *will* or *should* expectations, but not both, does not hamper our ability to obtain consistent estimates of α and γ.

same person responds to the analogous *will* and *should* expectations question relative to the respective mean response. These equations result in within-individual analyses. In particular, utilizing the multiple (repeated) measures aspect of our design enables us to control for all individual-specific factors that remain unchanged at time t. In our case, this is the actual service delivered to a given individual, the individual's "history" and characteristics, including a person's proclivity to use a specific portion of the response scale, and any new information received prior to time t. As a result, equations 13 and 14 no longer contain any unobserved variables and consistent estimation of two coefficients, $(\alpha_{jt} + \gamma_{jt})/(\alpha_{jt})$ and $(\alpha_{jt} + \gamma_{jt})$, is possible. From these two coefficients we can fully identify the two key structural parameters found in equations 1S and 3S. (Dividing $(\alpha_{jt} + \gamma_{jt})$ by $(\alpha_{jt} + \gamma_{jt})/(\alpha_{jt})$ yields an estimate of α_{jt}. Once $\hat{\alpha}_{jt}$ is obtained, it is easy to get an estimate of γ_{jt}.) These estimates can potentially falsify our original structural equations—if $\hat{\alpha}_{jt}$ is not significantly greater than zero and $\hat{\gamma}_{jt}$ is not significantly less than zero, there is strong evidence to disconfirm our conceptualization as stated in 1S, 2S, and 3S.

Results

We begin by reporting the estimates for the behavioral intention equation, 5S. Similar to our study 1 results in which perceptions of quality relate positively to behavioral intentions, we find that overall perceived quality positively and significantly ($t_{146} = 2.18$) relates to the index of behavioral intentions. We next explore the relationship between overall perceived quality and the measures of the five dimensions of service quality as posited by Parasuraman, Zeithaml, and Berry (1985). We start with the unconstrained form of this relationship, equation 7, which allows for different parameter values on each of the perception and expectation measures. Next, we estimate the gaps model, equation 6, which constrains the coefficient on the j^{th} dimension of expectations to equal the negative of the coefficient on the j^{th} dimension of perceptions. Finally, we estimate our dynamic model specification, equation 4S, which imposes the constraint that the coefficients on the expectation dimensions equal zero.[17]

We report the results of these estimates in Table 6. Column 1 of Table 6 corresponds to equation 7 and columns 2 and 3 correspond to equations 6 and 4S, respectively. We note all three equations yield results consistent with the hypothesis that the particular model being tested is statistically significant. Hence, we next explore which equation best captures reality by noting that equations 4S and 6 are constrained versions of 7, thereby enabling us to test the implied constraints of 4S and 6. We do so in the model comparison tests reported in the footnotes of Table 6. Specifically, the F-tests indicate that we must reject the constraint $\Theta_{1j} = -\Theta_{2j}$, but that we cannot reject the constraint $\Theta_{2j} = 0$. More generally, we reject the static formulation of the gaps model (i.e., equation 6) and its implied constraint. However, we fail to reject our dynamic model in favor of what is effectively an unconstrained version of the static gaps formulation (i.e., equation 7). We take these results to demonstrate strong support for this part of our dynamic specification.

Finally, we use the estimates in column 4 of Table 6 to test for the relevance of all five proposed dimensions of service quality. In this model, we eliminate all but the reliability and empathy perception variables. Comparing the estimates in column 4 with those in column 3 by means of an F-test indicates that we fail to reject the two-dimensional representation of quality in favor of

Table 6
OVERALL PERCEIVED QUALITY OF SERVICE EQUATION ESTIMATES

Independent variables	(1) Unconstrained model	(2) Gaps model	(3) Dynamic model	(4) Limited dynamic model
Gaps				
Reliability		.026*		
Responsiveness		−.003		
Assurance		−.006		
Empathy		.026**		
Tangibles		−.000		
Perceptions				
Reliability	.046*		.049*	.043*
Responsiveness	−.007		−.004	
Assurance	−.007		−.008	
Empathy	.027**		.021***	.015***
Tangibles	.003		.003	
Expectations[a]				
Reliability	−.004			
Responsiveness	.008			
Assurance	.005			
Empathy	−.022			
Tangibles				
R^2	.286	.214[b]	.272[c]	.266[d]

[a]We test whether the *will* and *should* expectation variables require different coefficients in this analysis. They do not, unsurprisingly, for two reasons. First, at any given time t, the *will* and *should* expectations differ only by an individual-specific constant. Second, the coefficient on the contemporaneous expectation is zero, whether for *will* or *should* expectations.

[b]Significantly different from unconstrained model at the .05 level.

[c]Not significantly different from unconstrained model.

[d]Not significantly different from unconstrained model or dynamic model.

*Significant at the .01 level.

**Significant at the .05 level.

***Significant at the .10 level.

[17]As the five dimensions of perceived service quality appear on the left side in our structural model, we tested for the necessity of two-stage (simultaneous) estimation of our overall service quality equation. This test (Hausman 1978) revealed a recursive relationship, indicating the appropriateness of ordinary least squares estimation. This was true for the behavioral intention equations as well.

Table 7
RELIABILITY AND EMPATHY PERCEPTION EQUATION ESTIMATES[a]

Dependent variable	*Will expectation coefficient* (α_j)[b]	*Should expectation coefficient* (γ_j)
Reliability perception	.771*	−.513*
	(.211)	(.218)
Empathy perception	.714*	−.372*
	(.115)	(.125)

[a]Standard errors are in parentheses. A technical appendix is available from the authors upon request explaining how standard errors and significance levels were calculated. The basic idea was to use Monté Carlo techniques to calculate the distribution of $(\alpha_j + \gamma_j)/[(\alpha_j + \gamma_j)/(\alpha_j)] = \hat{\alpha}$, and $(\alpha_j + \gamma_j) - \hat{\alpha} = \hat{\gamma}_j$ and then calculate the standard deviations and fractiles of these derived distributions.

[b]Because the coefficient on delivered service equals one minus α, the implied delivered service coefficients for reliability and empathy are .229 and .286, respectively.

*Significant at the .01 level.

the five-dimensional representation.[18] Thus, column 4 represents the preferred model for overall perceived quality for our particular application. As expected, these estimates indicate that reliability is the primary driver of overall quality perceptions.

We next turn our attention to estimating the updating parameters α_{jt} and γ_{jt} via our reduced-form equations 13 and 14 for the two relevant (i.e., significant) quality perception dimensions. Before discussing these estimates, however, we test whether the perceptions obtained from respondents providing *will* expectations differ from those of respondents providing *should* expectations by running regressions where perceptions are a function of the type of expectation measured. As we fail to find a significant coefficient on the version of the expectation measure variable for any of the perception dimensions, we infer that all of our perception data come from the same overall population.

Table 7 presents the results of our estimates of α_{jt} and γ_{jt}. First, we find that for both the reliability and empathy dimensions the estimate of the *will* expectation coefficient (i.e., α_{jt}) is significantly greater than zero but less than one, as postulated.[19] Second, our two estimates of the *should* expectation coefficient (i.e., γ_{jt}) are significantly less than zero, also as postulated. Thus, our field study results are compatible with our conceptualization that prior expectations of what service *will* occur positively influence perceptions of delivered service, whereas prior expectations of what service *should* occur negatively influence these perceptions. In addition, because α_{jt} also appears in equation 1S, we find support for our premise that *will* expectations are updated after a service encounter.

[18]We also fail to reject the parsimonious model given in column 4 in favor of the full unconstrained model given in column 1.

[19]Table 7 discusses how we develop significance tests.

DISCUSSION

We present a process model of how individuals develop perceptions of a firm's service delivery system over time. By explicitly acknowledging that perceptions and expectations change over time, we are better able to explicate and test the relationships between expectations, perceptions, and intended behavior. The model is tested with data derived from two very different studies, one a longitudinal laboratory experiment and the other a field study using questionnaire data collected at one point in time. In both cases, the results are strongly compatible with all aspects of our process model.

We find the convergence of results for the two different studies very encouraging. Our model appears robust to different analytic approaches, different data collection methods, and different service settings. Thus, though one might generate specific criticisms of the individual studies, we can think of none that span both studies. Consequently, we have a strong posterior belief that our model adequately summarizes the major forces that cause customers to form and update their perceptions of a firm's overall service quality level.

These forces have major implications for any firm interested in service quality. As expected, but never empirically verified in a field setting, our results indicate that the greater customers' perceptions of a firm's overall service quality, the more likely the customers are to engage in behaviors beneficial to the strategic health of the firm (e.g., generate positive word of mouth, recommend the service, etc.).

Our research also provides insights into how firms can best increase customers' perceptions of their overall service quality. Our most important managerial insight relates to the role of expectations. The prevailing model of service quality defines perceived service quality as the gap between expectations and perceptions, and does not differentiate among types of expectations. It leads to the strategic implication that firms can try either to increase perceptions or lower expectations in their quest to increase overall service quality. Our results are incompatible with both this one-dimensional view of expectations and the gap formulation for service quality. Instead, we find that service quality is directly influenced only by perceptions. Also, increasing customer expectations of what a firm *will* provide during future service encounters actually leads to higher perceptions of quality after the customer is exposed to the actual service, all else equal. From this finding we infer that firms should manage customers' predictive expectations *up* rather than down if they want to increase customer perceptions of overall service quality. In addition, our results strongly support our premise that customers' expectations of what a firm *should* deliver during a service encounter *decrease* their ultimate perceptions of the actual service delivered, all else equal. Therefore, improved assessments of service quality can result when customers' expectations of what a firm *should* deliver are managed downward.

The issue of managerial importance, then, is how to manage both types of expectations. Ideally, one would want to simultaneously increase customers' *will* expectations and decrease their *should* expectations. At this stage of our research, we know of no activity that can ensure this result. One airline firm attempted to do this by simultaneously telling customers that all airlines had problems with guaranteeing on-time arrivals because of factors outside the airlines' control, but that they were the best at being on time. In this way the firm's ad campaign attempted to address both the *should* and *will* expectations. Whether or not this approach to managing both sets of expectations worked as intended is an empirical question.

A second approach to managing *will* and *should* expectations is for the firm to engage in activities that increase the customers' *will* expectations without a proportional increase in their expectations of what the firm *should* do. From equations 1S and 2S, we see that providing the best possible service each and every time can increase *will* expectations but it might also increase the *should* expectations. Fortunately, our empirical evidence suggests that the *will* expectations increase faster than the *should* expectations, so that the net impact on perceptions is positive. However, firms need to monitor the relative magnitudes of α, β, and γ to ensure that increases in objective service quality also result in increases in perceptions of service quality (see footnote 13 for more details). Finally, managers may be able to identify specific firm actions (other than service) that affect only the *will* or *should* expectations. Such actions would enable the firm to increase (decrease) the *will* (*should*) expectations without modifying the other.

In addition to providing managerial insights, we were able to demonstrate a method of estimating the two key parameters from our dynamic model by using survey data taken from customers at only one point in time. As a result, managers can learn about the relative importance of service delivery and customer expectations for their specific business. This determination should be very useful in assessing the relative value of trying to modify perceptions through changes in the service delivery system and the firm's communications, as well as identifying the speed with which managers can expect perceptions to change over time.

We believe our analytic approach provides managers an easily implementable method for estimating our model because it does not require measuring the actual service provided or prior expectations. However, as seen from our derivation, the estimation technique requires that (1) the surveys obtain multiple measures of perceptions and expectations, (2) all of the measures within a dimension have identical influence on that dimension, and (3) if the managers believe customers have much different levels of prior experience, they segment the customers so as to reflect the possible differences in the updating parameters.

Our research also has implications for academicians. We note a great similarity between our work on modeling perceived service quality and its impact on intended future behavior and the models of Churchill and Surprenant (1982) and Tse and Wilton (1988), who were concerned with explicating the factors that influence perceived product performance (and ultimately its impact on consumer satisfaction). As in our study 1, both of these research teams were able to measure prior expectations and the actual product performance. However, only Tse and Wilton measured two types of expectations and thus were able to obtain unbiased estimates. Their study found, analogous to our results, that prior *will* expectations and actual product performance were positively related to perceived performance. In addition, they found that prior *expectations on what consumers would ideally like to see* in the product were negatively related to perceived performance. Interestingly, they found the actual product performance variable to have a much stronger influence on perceived performance than we did in our study. This difference is not surprising given that services typically have a higher proportion of experience and credence properties than products, making service performance more difficult to evaluate than product performance. It seems likely that perceptions will be more influenced by expectations (relative to actual service) for firms with a higher content of unobservable (or fallible) quality. Along these lines, future research might assess the degree to which different industries or customers with different levels of prior experience influence the extent to which prior knowledge, new communications, or the actual service encounter dominates the process by which customers form judgments of quality.

Though we suggest conceptually, and demonstrate empirically, that customers update their expectations and perceptions, interesting aspects of this process have not been investigated. For example, the antecedents of the different expectation variables remain largely unexplored.[20] Given the need to manage *will* expectations up and the *should* expectations down, understanding the determinants of these expectations is a critical managerial issue. Also, because we can restate our equations mathematically in a variety of formats, our empirical analyses provide no evidence on the cognitive process by which customers form, store, or retrieve perceptions. Consequently, we hope that researchers utilize experimental and panel data to continue delving into the dynamic process by which customers form expectations and perceptions of service quality.

Finally, we note that our process model has the potential for broader applications. First, one might view overall service quality as a measure of the firm's service equity. Further, because the antecedents of this construct are known, measuring and managing these antecedents

[20]To date we know that word-of-mouth communications and information from expert sources affect *will* expectations whereas information on the competitors and to a lesser degree word of mouth affect *should* expectations. We used this knowledge to manipulate the subjects' prior expectations in study 1.

(e.g., expectations) can help a firm better understand which actions either enhance or detract from the firm's service equity and thus its ability to compete. Second, we see no reason why our process model would not apply to products as well as services. However, empirical support for this belief remains to be provided. Third, we see direct applicability of our model in better understanding, tracking, and influencing customer satisfaction as referred to in the popular press. The reason is that the measures used to reflect satisfaction are usually cumulative, versus transaction specific, and thus are analogous to our construct of perceived quality.

APPENDIX A
SCALING CALIBRATION

Our goal is to rescale all of the measures so that they have a common metric. Assume this metric is defined in terms of the measured expectations, *MWE* and *MSE*. For our measure of the delivered service construct, we note that only two levels of service, high and low, were experienced by the subjects. Let $MDS_{it} = 1$ if the service was high and $MDS_{it} = 0$ if the service was low, where MDS_{it} equals measured delivered service. Next, we define

$$DS^*_{it} = a + b\,MDS_{it}, \quad \text{(A1)}$$

where a and b are ≥ 0. This formulation enables us to convert our measure of delivered service onto the same 100-point scale as the expectation scale. Thus, the a parameter represents the metric value of low service, whereas $a + b$ represents the metric value of high service.

We also need to acknowledge that a person's measured perception, denoted *MPS*, may be on a different scale than the person's measured expectations. For calibration across these scales, we define

$$PS_{it} = c + MPS_{it}, \quad \text{(A2)}$$

where c is a shift parameter.

Similarly, we let the measured behavioral intention, denoted *MBI*, be on a different scale than the person's measured perceptions:

$$BI_{it} = d + MBI_{it}, \quad \text{(A3)}$$

where d is again a shift parameter.

APPENDIX B
DERIVATION OF ESTIMATING EQUATIONS

First, using equation 12, write equation 3S completely in terms of either the *will* or *should* expectation:

$$PS_{ijt} = (\alpha_{jt} + \gamma_{jt})WE_{ijt-1} + (1 - \alpha_{jt})DS^*_{ijt} + \epsilon_{B1it}, \quad \text{(B1)}$$

and

$$PS_{ijt} = (\alpha_{jt} + \gamma_{jt})SE_{ijt-1} + (1 - \alpha_{jt})DS^*_{ijt} + \epsilon_{B2it}. \quad \text{(B2)}$$

These equations make it clear that even if measures of lagged expectations are available, using only one of the two lagged expectation variables to estimate the relationships between expectations and perceptions (even after controlling for actual service) will result in biased expectation coefficient estimates for that expectation (i.e., the obtained estimate is $(\alpha \hat{+} \gamma)$ versus $\hat{\alpha}$ or $\hat{\gamma}$).[21]

Next, because we only observe *contemporaneous* expectations, we use equation 1S or 2S to rewrite equations B1 and B2 in terms of current values of either *WE* or *SE*.

$$PS_{ijt} = \left(\frac{\alpha_{jt} + \gamma_{jt}}{\alpha_{jt}}\right) WE_{ijt} - \gamma_{jt}\frac{(1 - \alpha_{jt})}{\alpha_{jt}} DS^*_{ijt} + \epsilon_{B3it}, \quad \text{(B3)}$$

and

$$PS_{ijt} = (\alpha_{jt} + \gamma_{jt})\,SE_{ijt} + [1 - \alpha_{jt}\beta_{jt}K_{ijt} - \gamma_{jt}\beta_{jt}K_{ijt} - \alpha_{jt}]DS^*_{ijt} + \epsilon_{B4it}. \quad \text{(B4)}$$[22]

Next, imagine that equations B3 and B4 have n subscripts indicating the individual measures for the perception and expectation constructs. Take means over the n items for the i^{th} individual and the j^{th} dimension in equations B3 and B4, yielding

$$PS_{ij\cdot t} = \left(\frac{\alpha_{jt} + \gamma_{jt}}{\alpha_{jt}}\right) WE_{ij\cdot t} + \epsilon_{B5it}, \quad \text{(B5)}$$

and

$$PS_{ij\cdot t} = (\alpha_{jt} + \gamma_{jt})\,SE_{ij\cdot t} + \epsilon_{B6it}, \quad \text{(B6)}$$

where the $\cdot$ notation indicates the mean for the i^{th} individual on the j^{th} dimension. Subtracting equations B5 and B6, respectively, from equations B3 and B4 supplemented with the n subscripts produces equations 13 and 14 reported in the text.[23]

REFERENCES

Amemiya, Takeshi (1985), *Advanced Econometrics*. Cambridge, MA: Harvard University Press.

Anderson, Eugene W. and Mary W. Sullivan (1990), "Customer Satisfaction and Retention Across Firms," presentation at the TIMS College of Marketing Special Interest Conference on Services Marketing, Nashville, TN (September).

Bearden, William D. and Jesse E. Teel (1983), "Selected Determinants of Customer Satisfaction and Complaint Reports," *Journal of Marketing Research*, 20 (November), 21–8.

Bitner, Mary Jo (1990), "Evaluating Service Encounters: The Effects of Physical Surroundings and Employee Responses," *Journal of Marketing*, 54 (April), 69–82.

Bolton, Ruth N. and James H. Drew (1991a), "A Multi-Stage Model of Customers' Assessments of Service Quality and Value," *Journal of Consumer Research*, 17 (March), 375–84.

---

[21]Most published studies linking expectations to perceptions include only one expectation in the estimating equation. If our laboratory results generalize, then equations B1 and B2 imply the obtained coefficients in these single-expectation models are biased.

[22]Note that these equations imply that use of only the contemporaneous *will* or *should* expectation in a perceptions equation results in biased estimates for these expectation variables.

[23]If one measures both *will* and *should* expectations, one can derive a third estimating equation containing both types of expectations (see Boulding et al. 1992).

——— and ——— (1991b), "A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes," *Journal of Marketing*, 55 (January), 1-9.

Boulding, William, Richard Staelin, Ajay Kalra, and Valarie Zeithaml (1992), "Conceptualizing and Testing a Dynamic Process Model of Service Quality," Working Paper No. 92-121. Cambridge, MA: Marketing Science Institute.

Cadotte, Ernest R., Robert B. Woodruff, and Roger L. Jenkins (1987), "Expectations and Norms in Models of Consumer Satisfaction," *Journal of Marketing Research*, 24 (August), 305–14.

Churchill, Gilbert A., Jr. (1979), "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, 11 (August), 254–60.

——— and Carol Surprenant (1982), "An Investigation Into the Determinants of Satisfaction Research," *Journal of Marketing Research*, 19 (November), 491–504.

Day, Ralph L. (1977), "Towards a Process Model of Consumer Satisfaction," in *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*, H. Keith Hunt, ed. Cambridge, MA: Marketing Science Institute, 153–83.

Forbes, J. D., David K. Tse, and Shirley Taylor (1986), "Toward a Model of Consumer Post-Choice Response Behavior," in *Advances in Consumer Research*, Vol. 13, Richard L. Lutz, ed. Ann Arbor, MI: Association for Consumer Research, 658–61.

Garvin, David A. (1987), "Competing on the Eight Dimensions of Quality," *Harvard Business Review*, 65 (November–December), 101–9.

Gilly, Mary C. (1979), "Complaining Consumers: Their Satisfaction With Organizational Response," in *New Dimensions of Consumer Satisfaction and Complaining Behavior*, Ralph L. Day and H. Keith Hunt, eds. Bloomington, IN: School of Business, Indiana University, 99–107.

———, William L. Cron, and Thomas E. Barry (1983), "The Expectation-Performance Comparison Process: An Investigation of Expectation Types," in *International Fare in Consumer Satisfaction and Complaining Behavior*, Ralph L. Day and H. Keith Hunt, eds. Bloomington, IN: School of Business, Indiana University, 10–16.

Hausman, Jerry A. (1978), "Specification Tests in Econometrics," *Econometrica*, 46 (November), 1251–72.

Hjorth-Anderson, Chr. (1984), "The Concept of Quality and the Efficiency of Markets for Consumer Products," *Journal of Consumer Research*, 11 (2), 708–18.

Holbrook, Morris B. and Kim P. Corfman (1985), "Quality and Value in the Consumption Experience: Phaedrus Rides Again," in *Perceived Quality*, J. Jacoby and J. Olson, eds. Lexington, MA: Lexington Books, 31–57.

Kalra, Ajay (1992), "An Empirical Validation and Transaction-Level Investigation of an Expectations-Based Process Model of Service Quality," PhD dissertation, Duke University.

LaTour, Stephen A. and Nancy C. Peat (1979), "Conceptual and Methodological Issues in Consumer Satisfaction Research," in *Advances in Consumer Research*, Vol. 6, William L. Wilkie, ed. Ann Arbor, MI: Association for Consumer Research, 431–7.

Lynch, John G., Jr., Dipankar Chakravarti, and Anusree Mitra (1991), "Contrast Effects in Consumer Judgments: Changes in Mental Representations or in the Anchoring of Rating Scales?" *Journal of Consumer Research*, 18 (3), 284–97.

Maynes, E. Scott (1976), "The Concept and Measurement of Product Quality," *Household Production and Consumption*, 40 (5), 529–59.

Miller, John A. (1977), "Studying Satisfaction, Modifying Models, Eliciting Expectations, Posing Problems, and Making Meaningful Measurements," in *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*, H. Keith Hunt, ed. Bloomington, IN: School of Business, Indiana University, 72–91.

Oliver, Richard L. (1977), "Effect of Expectation and Disconfirmation on Post-Exposure Product Evaluation: An Alternative Interpretation," *Journal of Applied Psychology*, 62 (April), 480–6.

——— (1980), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, 17 (November), 460–9.

Olson, Jerry C. and Philip Dover (1979), "Disconfirmation of Consumer Expectations Through Product Trial," *Journal of Applied Psychology*, 64 (April), 179–89.

Parasuraman, A., Leonard L. Berry, and Valarie A. Zeithaml (1990), "An Empirical Examination of Relationships in an Extended Service Quality Model," working paper, Marketing Science Institute.

———, Valarie A. Zeithaml, and Leonard L. Berry (1985), "A Conceptual Model of Service Quality and Its Implications for Future Research," *Journal of Marketing*, 49 (Fall), 41–50.

———, ———, and ——— (1988), "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality," *Journal of Retailing*, 64 (Spring), 12–40.

Prakash, Ved (1984), "Validity and Reliability of the Confirmation of Expectations Paradigm as a Determinant of Consumer Satisfaction," *Journal of the Academy of Marketing Science*, 12 (Fall), 63–76.

Swan, John E. and Frederick I. Trawick (1980), "Satisfaction Related to Predictive vs. Desired Expectations: A Field Study," in *New Findings on Consumer Satisfaction and Complaining*, Ralph L. Day and H. Keith Hunt, eds. Bloomington, IN: School of Business, Indiana University, 15–22.

——— and ——— (1981), "Disconfirmation of Expectations and Satisfaction With a Retail Service," *Journal of Retailing*, 57 (Fall), 49–67.

Tse, David K. and Peter C. Wilton (1988), "Models of Consumer Satisfaction Formation: An Extension," *Journal of Marketing Research*, 25 (May), 204–12.

Wilton, Peter C. and M. Nicosia (1986), "Emerging Paradigms for the Study of Consumer Satisfaction," *European Research*, 14 (January), 4–11.

Woodruff, Robert B., Ernest R. Cadotte, and Roger L. Jenkins (1983), "Modeling Consumer Satisfaction Processes Using Experience-Based Norms," *Journal of Marketing Research*, 20 (August), 296–304.

Zeithaml, Valarie A. (1988), "Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence," *Journal of Marketing*, 52 (July), 2–22.

———, Leonard L. Berry, and A. Parasuraman (1991), "The Nature and Determinants of Customer Expectations of Service," working paper, Marketing Science Institute.

Reprint No. JMR301101