Distinguishing Service Quality and Customer Satisfaction: The Voice of the Consumer

Dawn Iacobucci and Amy Ostrom
Department of Marketing
Kellogg Graduate School of Management
Northwestern University

Kent Grayson
Department of Marketing
London Business School

Service quality and customer satisfaction are important concepts to academic researchers studying consumer evaluations and to practitioners as a means of creating competitive advantages and customer loyalty. This article presents two studies that rely on divergent methodologies to examine whether or not quality and satisfaction have distinct antecedent causes, consequential effects, or both (i.e., whether or not they should be considered a single construct, or distinct, separable constructs). We focus on consumers' understanding and use of the words quality and satisfaction; in both studies, respondents report whether or not they think quality and satisfaction differ, and if so, on what dimensions or under what circumstances. In the first study, we use the qualitative “critical incident” technique to elicit service attributes that are salient to respondents when prompted to consider quality and satisfaction as distinct. We code the responses to these open-ended survey questions to examine whether quality can be teased apart from satisfaction, from the respondents’ (consumers’) perspective. In the second study, to triangulate on the qualitative data, we experimentally manipulated a number of service attributes drawn from both the first study and from the literature to see whether or not they have differential impacts on judgments of quality and satisfaction. We did not presuppose that quality and satisfaction differ—rather, we asked respondents to make a judgment either of quality or of satisfaction, defining the term as they saw fit. We inferred from their judgments.

Requests for reprints should be sent to Dawn Iacobucci, Department of Marketing, Kellogg Graduate School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208.
whether the terms were used differently or interchangeably. The results of the two studies offer fairly robust consumer definitions of quality and satisfaction.

Many industries are paying greater attention to service quality and customer satisfaction, for reasons such as increased competition and deregulation (Reichheld & Sasser, 1990; Schlesinger & Heskett, 1991). Academics have also been studying quality and satisfaction to understand determinants and processes of customer evaluations (Bitner & Hubbert, 1993; Boulding, Staelin, Kalra, & Zeithaml, 1993; Cadotte, Woodruff, & Jenkins, 1987; Churchill & Surprenant, 1982; Fornell, 1992; Oliver, 1993; Parasuraman, Berry, & Zeithaml, 1985; Westbrook & Oliver, 1991).

The academic literature postulates that customer satisfaction is a function of the discrepancy between a consumer's prior expectations and his or her perception regarding the purchase (Churchill & Surprenant, 1982; Oliver, 1977; Tse & Wilton, 1988; Yi, 1990). When an experience is better than the customer expected, there is thought to be positive disconfirmation of the expectation, and a favorable customer evaluation is predicted. Service quality is defined similarly, as a comparative function between consumer expectations and actual service performance (Parasuraman et al., 1985). In the customer satisfaction literature, this model is referred to as the "Disconfirmation Paradigm"; in the service quality literature, it is referred to as the "Gap Model."

Sometimes the terms quality and satisfaction are used interchangeably (both in industry and academia), as if the two are essentially one evaluative construct. Given the aforementioned definitions, the two appear highly similar. However, several researchers are interested in how they differ (cf. Dabholkar, 1993; Gotlieb, Grewal, & Brown, 1994). For example, some service quality researchers describe satisfaction as a more specific, short-term evaluation (e.g., evaluating a single service encounter) and quality as a more general and

---

1Churchill and Surprenant (1982) said that "the vast majority of [customer satisfaction] studies have used some variant of the disconfirmation paradigm" (p. 491); Oliver (1993) concluded that a variety of authors' definitions of satisfaction are "consistent with the expectancy disconfirmation model" (p. 72); and Tse and Wilton (1988) said that "it is generally agreed that [satisfaction] can be defined as . . . the evaluation of the perceived discrepancy between prior expectations . . . and the actual performance of the product" (p. 204).

2Parasuraman et al. (1985) stated that "[service] quality is a comparison between expectations and performance" (p. 42) and recently reiterated service quality as "the discrepancy between customers' expectations and perceptions" (Parasuraman, Zeithaml, & Berry, 1994, p. 111); Boulding et al. (1993) modeled perceptions of quality as a function of "prior expectations of what will and what should transpire . . . and the actual delivered service" (p. 7), although Cronin and Taylor (1994) found no support for expectations, but rather found that perceptions of quality are only a function of perceived product performance and stated that the findings of Boulding et al. concur. Finally, Teas (1994) explored the comparison between expectations and perceived performance in a thoughtful analysis of mathematical functional forms and preferred ideal points to expectation norms.
long-term evaluation (cf. Bitner & Hubbert, 1993; Parasuraman et al., 1985). In contrast, some customer satisfaction researchers posit quality as the more specific judgment and a component of satisfaction, the broader evaluation (cf. Oliver, 1993). If these concepts are distinct, then they are worthy of further separate pursuit, but if they are the same, then more efficient theoretical progress would be made if these concepts were studied via convergence in a shared literature.

The majority of articles attempting to distinguish quality and satisfaction have been conceptual in nature (cf. Oliver, 1993). Further, much of this conceptualization has been driven by the perspective of the researcher. We too have opinions regarding how quality and satisfaction might be most suitably defined, but our opinions, and those stated thus far in the literature, are infinitely less important than empirical support demonstrating the viability of a set of hypothesized definitions and relations. Notwithstanding the importance of theoretical distinctions, a critical issue for both researchers and marketing managers is whether or not consumers also see such distinctions. If consumers prompted to evaluate either "quality" or "satisfaction" respond with some common, omnibus evaluation and the evaluations converge, then we know something about the consumer's lack of differentiation. Further, if consumers treat quality and satisfaction as one concept, but academicians treat them as two, then the latters' distinctions would not be testable unless measurement and data collection procedures carefully probed and elicited the desired differences.

In this article, we first discuss what evidence must be sought in order to demonstrate whether quality and satisfaction are distinct. We then present two empirical studies. The first is an analysis of open-ended survey questions in which we ask participants to consider how quality and satisfaction differ. The second is a series of experiments in which we ask participants to make judgments of quality or satisfaction, using respondent-determined definitions. In both studies, quality and satisfaction are examined simultaneously to understand which purchase attributes may serve as differential antecedents. In addition, participants are allowed to determine the meanings of the words quality and satisfaction. We rely on divergent methodologies (content analysis of

---

3This might be partly due to the difficulty of testing some of the proposed distinctions. For example, the aforementioned distinction of specific, short-term versus global, long-term evaluations carries an inherent part-whole relation (which can be difficult to detect), as does the conceptualization of quality as cognitive and satisfaction as both cognitive and affective (cf. Cadotte et al., 1987; Mano & Oliver, 1993; Oliver, 1993; Westbrook & Oliver, 1991). Other hypothetical distinctions might require data such as financial indicators (cf. Bearden & Teel, 1983; Fornell, 1992; Reichheld & Sasser, 1990) that are difficult to obtain due to proprietary issues. Nevertheless, certain other features might be more readily tested, and we focus on them in this article.

4In this sense, our research is like that of Zeithaml (1988), which elicited consumers' various uses of the term value.
qualitative data and experimentation) in order to triangulate on these research questions and thereby enhance the validity of the results.

ESTABLISHING EVIDENCE FOR ONE VERSUS TWO CONSTRUCTS

In this section, we discuss briefly how it may be determined whether two concepts are separable and unique or are better conceptualized as one construct. Three related logics regarding constructs' uniquenesses are described: (a) constructs' unique positions in nomological networks, (b) conceptual orthogonality, and (c) uniqueness via structural equations modeling.

Marketers are familiar with nomological nets as theoretical representations of hypothetical constructs and their interrelations (cf. Bagozzi, Yi, & Phillips, 1991; Cronbach & Meehl, 1955). Concepts are separable theoretical constructs if they occupy unique positions in a nomological network as determined by unique sets of antecedent causes, consequential effects, or both (cf. Sternthal, Tybout, & Calder, 1987). Conversely, if two network concepts share all theoretical antecedents and consequences, then they are "structurally equivalent," or logically isomorphic, and to discuss them as if they were unique would be indefensible and empirically untestable. For example, Figure 1 represents the standard definitions of quality and satisfaction that share antecedents (expectations and perceptions of the purchase experience) and common consequences (e.g., repeat purchase intentions). The positions of quality and satisfaction in this nomological network are not unique, but structurally interchangeable. Other factors differentially causing or affecting quality and satisfaction must be both conceptualized and tested if the two are to be distinguished. We attempt to develop and test the beginnings of such a network in this article.

Another means of thinking about unique causes and effects is to consider whether or not two constructs may be conceptualized as orthogonal. If two concepts shared all causes, they could not vary independently. Thus, quality and satisfaction can be distinguished if one can hypothesize circumstances for which, say, a high-quality product can result in customer satisfaction or dissatisfaction. Nevertheless, empirical tasks may still be challenging. Given mea-

![Figure 1: Consumer evaluation judgments: Service quality and customer satisfaction.](image-url)
surement error, constructs that are orthogonal (theoretically distinct) or coincidental (theoretically identical) will often both yield measured indicators that are "somewhat" correlated (Iacobucci, Grayson, & Ostrom, 1994). Indeed, differences between constructs in nomological nets (in their antecedents, consequences, or both) need not be strictly qualitative in nature; quantitative differences in degree or in functional form are logically sound (Katz, 1962); however, such differences are more difficult to demonstrate empirically than differences in kind.

Many researchers have commented on the correlations among measures of quality and satisfaction (e.g., Bearden & Teel, 1983; Cadotte et al., 1987; Tse & Wilton, 1988), and structural equations models (e.g., Bagozzi et al., 1991; Gerbing & Anderson, 1988) have been proposed as a method for providing empirical evidence for construct validity. However, recent developments have shown that the method (like any other) is not without its limitations (Lee & Hershberger, 1990; MacCallum, Wegener, Uchino, & Fabrigar, 1993). Many researchers would argue that causation is best studied via experimentation (Iacobucci, 1994), and we do so in the studies we report shortly. Although experimentation has been used in some satisfaction studies, it has not been used in the service quality literature or in studies that examine quality and satisfaction simultaneously.

To illustrate a limitation of structural equations modeling, consider the following. Some researchers are interested in the direction of causality between quality and satisfaction, but in a network such as that depicted in Figure 1, the directionality is unknowable. Quality and satisfaction are structurally equivalent in this network, so applying this model to any arbitrary data set would yield identical fit statistics whether quality was hypothesized as affecting satisfaction, or the reverse, or the path was made bidirectional (Lee & Hershberger, 1990; MacCallum et al., 1993). In fact, it is likely premature to query the causal direction between quality and satisfaction because such a question presumes distinction between the two. Uniqueness must first be established. (If identity is established instead, the subsequent question would be invalid.) Moreover, not only is it important to identify unique antecedents, consequences, or both, theoretically, one must also do so empirically (the mathematical tool of structural equations, or any other, cannot overcome the logical impossibility of teasing apart structurally equivalent concepts). In summary, given the conceptualization in Figure 1, it would be more precise and parsimonious if the two nodes were collapsed to one.

If consumers essentially think that quality and satisfaction are only different operationalizations of one construct, say "consumer evaluation," the resulting

---

3Bolton and Drew (1991) proposed a model identical to Figure 1 in which quality, satisfaction, and purchase intentions are replaced with satisfaction, disconfirmation, and quality. Their hypothesized effect of disconfirmation on satisfaction is therefore unknowable.
advantage is conceptual simplicity. This parsimonious stance would also allow
greater leverage across the two literatures: Variables could be viewed as repli-
cates, and scientific progress could be advanced more quickly. However, if
consumers consider quality and satisfaction to be two constructs, the resulting
advantage would be the required richness in the hypothesized connections in
their joint nomological net. Support would require different effects on or effects
of quality and satisfaction, or both; thus, if quality and satisfaction are to be
considered separate, though likely related constructs, the next logical question
to address would be, "How are they different?"

EMPIRICAL STUDIES

To begin to address the question of whether or not quality and satisfaction
might differ, we conducted two studies. Both studies share a property we
believe to be very important for this article—a focus on the consumer voice:
We asked participants to respond using their own interpretations of the terms
quality and satisfaction. We can make much ado about differences between
these concepts in the literature, but if in mental representation or in response
to measurement probes of satisfaction or quality a consumer simply interprets
either phrase interchangeably as an overall evaluation, "how good was it,"
then the analysis in the literature would correspond poorly to consumers’
processing of the same terms. Furthermore, the possibility exists that consum-
ers may conceive of the constructs as different in ways that the conceptualiza-
tions had not anticipated.6

The two studies differ in that the first has a more qualitative flavor than the
second. The first study relies on a standard services marketing research tool—
the critical incident method. The resulting qualitative data are coded for con-
sumer statements that might distinguish quality and satisfaction. The second
study used a set of factorial experiments in which manipulations are imple-
mented and participants make judgments along rating scales. Our goal in using
these different methods is triangulation on the quality and satisfaction con-
structs.

The studies also differ in that the first presumes that quality and satisfaction
are different constructs, and respondents are asked to describe the nature of the
differences. In the second study, we allow the respondents to indicate whether
they naturally (i.e., unprompted) consider quality and satisfaction to be the
same or different. Even finding a few differences in the studies that follow

---

6 Many researchers recognize the importance of conducting marketing research in ways that
allow the consumer voice to speak, rather than only respond to researcher-determined parame-
ters (Green, Wind, & Jain, 1973; Henderson, 1995; Steenkamp, Van Trijp, & Ten Berge, 1994;
would begin to establish uniqueness, but presumably a strong case for distinct constructs would rest on the discovery of multidimensional differences.

Study 1

Participants. Seventy-seven master of business administration students from a midwestern business school answered the questions described shortly. Of the 77 participants, 57% were male, their average age was 30, and their average work experience was 7 to 8 years.

Survey. We followed closely the methodology of Bitner, Booms, and Tetreault (1990) in eliciting critical incidents and that of Bowers (1970) in the content analysis of the resulting data. Bitner et al. (1990) defined critical incidents as specific interactions between customers and employees that are particularly good or bad. Respondents were asked to evaluate a recalled encounter in terms of quality and satisfaction (see Appendix A). Our study added an experimental factor: We asked some participants to describe a service encounter that "you would characterize as a high quality service, yet you were dissatisfied as a customer anyway" and others to describe an encounter that "you would describe as a low quality service, yet you were satisfied as a customer anyway" (38 participants described the former, 39 the latter).

This format presumes a distinction between the two constructs and requests that respondents recognize the difference. These instructions increase the likelihood that the two concepts will be described as different. If they are not, such results would be very conservative evidence that consumers, even when forced to do so, could not see the two concepts as different. However, if the two are described differently, then we will have some evidence regarding the dimensions and features on which respondents most readily distinguish quality from satisfaction. Note too that it allows for respondents' idiosyncratic use of the terms, and it will be our task to infer their meaning and possible differences.

We performed a content analysis (cf. Berelson, 1971; Bowers, 1970; McCracken, 1991) on items "d" and "e" in Appendix A, following the process of Bowers (1970). Content analyses share the advantages (and disadvantages) of qualitative methods, such as the capacity to interpret "people's accounts of events without depriving these accounts of their power or eloquence" (Viney, 1983, p. 560).

The data were divided into thought units, which were then classified using 18 codes, "A" through "R" (summarized in Appendix B). Whereas our categories were developed from the data themselves, the names for the resulting categories were drawn as much as possible from the literature. For example, when the theme of a thought unit resembled "responsiveness" or "empathy" (Parasuraman et al., 1985), or "physical aspects of the environment" (Bitner,
1992), and so forth, we labeled them as such to enhance the correspondence between the data and the literature (references are provided in Appendix B). Each thought unit was also rated for positive or negative valence; for example, if a respondent mentioned speed, was the service quick or slow; if a respondent mentioned friendliness, was the provider friendly or rude, and so forth.\footnote{Two judges divided the messages into 451 thought units, with agreement of 93%. After trial classification, the categories were refined, and four judges classified the units with conservative agreement (three or four of four raters) at 60%. When two or three raters agreed, the majority code was recorded. If two raters agreed on one code but two agreed on another, researchers resolved the conflict. If all raters disagreed, a code of "R" was recorded (uncodable). All judges were blind to the study.}

We explored three analyses of these data. In the first, a frequency was assessed of all mentions of each category (e.g., a participant responding with three instantiations of the code "friendliness" resulted in a score of 3). In the second analysis, we standardized so that more verbose respondents' data would not be weighted more: Multiple mentions of a category were analyzed as only one (i.e., the data were made binary; in the friendliness example, the score would be 1—friendliness was mentioned, and additional friendliness themes were interpreted as synonymous points expounding on the same property). Finally, in the third analysis, we took the ipsative norming further in correcting for "general" chattiness. For example, one respondent might yield 4 "friendliness" units, 5 "timeliness" units, 10 "expertise" units, and so on, whereas another respondent might have mentioned 2, 0, and 1 such codes. To ensure that each participant's data were equally weighted, we created proportions by dividing every frequency by the total number of thought units recorded for that respondent.

The transformations turned out to be noncritical because the results converged. In comparing the first and second treatments of the data, there was 98.1% agreement in terms of which codes were significant and which were not, and there was 97.2% agreement in the results between the first and third, and second and third treatments of the data. The results we will present will be those from the third transformation (percent codes given the total number of thought units). These data were proportions, so we also conducted an arc sine transformation (as per analyses of variance [ANOVAs] on proportions), but this transformation was also not critical, given that there was 98.1% agreement with the third set of results.

Table 1 contains mean proportions for those codes with significant effects. A value in Table 1 reflects the proportion of times the particular code appeared under the stated condition. These codes appeared (or their themes were stated) in different frequencies across conditions for satisfaction versus quality. The tests were the results of a mixed-design ANOVA, where the between-subjects factor was whether the respondent had been asked to report a high-quality,
TABLE 1
Mean Proportions From Study 1 for Significant Effects

<table>
<thead>
<tr>
<th>Category</th>
<th>Content</th>
<th>(d) Quality</th>
<th>(e) Satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. Positive tangible response to a mistake</td>
<td>HighQ, LowCS</td>
<td>.054</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>LowQ, HighCS</td>
<td>.004</td>
<td>.112</td>
</tr>
<tr>
<td>D. Negative recognizing customer value</td>
<td>HighQ, LowCS</td>
<td>.013</td>
<td>.120</td>
</tr>
<tr>
<td></td>
<td>LowQ, HighCS</td>
<td>.112</td>
<td>.034</td>
</tr>
<tr>
<td>H. Negative expertise</td>
<td>HighQ, LowCS</td>
<td>.000</td>
<td>.061</td>
</tr>
<tr>
<td></td>
<td>LowQ, HighCS</td>
<td>.191</td>
<td>.000</td>
</tr>
<tr>
<td>K. Positive physical aspects of environment</td>
<td>HighQ, LowCS</td>
<td>.038</td>
<td>.007</td>
</tr>
<tr>
<td></td>
<td>LowQ, HighCS</td>
<td>.013</td>
<td>.108</td>
</tr>
</tbody>
</table>

Note. Q = quality; CS = customer satisfaction.

The means in Table 1 are proportions, so ratio comparisons are meaningful (e.g., for code “B,” .112 is more than twice .054, etc.). The data on that code (capable tangible responsiveness of a service provider to a service mistake) suggest that service recovery is more important to (and therefore more clearly an antecedent of) judgments of satisfaction than quality.

The next code was for those encounters in which respondents felt service providers did not “recognize their value” as customers. The impact of this factor on quality and satisfaction is in the same direction and nearly equal in size (.013 vs. .112; .120 vs. .034). Not being recognized as a valuable customer negatively impacted both judgments of quality in low-quality situations and dissatisfaction in dissatisfactory situations. Thus, although this is clearly an important managerial factor (i.e., service providers must be trained to act like their customers have value), it is not a factor that distinguishes between quality and satisfaction.

Expertise had a greater impact on perceptions of quality than on satisfac-
tion. Incompetence fueled judgments of low quality (.191 vs. .000), but affected dissatisfaction less dramatically (.061 vs. .000). Finally, pleasant physical environments were more critical to satisfaction (.108 vs. .007) than to quality (.038 vs. .013). This finding is interesting because the importance of physical surroundings is thought to be a property that distinguishes customers' evaluations of services from that of goods (e.g., Bitner, 1992), and it is usually discussed in terms of service quality. This result suggests that physical environments are important to customers' evaluations even more broadly.

In summary, these qualitative data suggest three distinctions between quality and satisfaction. Service recovery and physical environments had a greater impact on satisfaction than on quality, and the expertise of service personnel had a stronger association with quality than with satisfaction. Thus the data show that when prompted to think about quality as distinct from satisfaction, participants can indeed do so. Moreover, the properties most salient to them in describing the nature of the distinction were service recovery, physical environments, and expertise. In the next study, we do not suggest a priori to participants that quality and satisfaction differ, but rather examine if they indicate any such differences unaided and if the responses resemble the results of the first study.

Study 2

Participants. Forty-three master of business administration students from a midwestern business school answered the questions described shortly. Sixty percent of the participants were male, their average age was 28, and their average work experience was 4 to 5 years. Participants were randomly assigned to the study conditions.

Survey. Each participant completed 10 stimulus–response judgments. The stimuli were brief scenarios that the respondents evaluated in terms of quality or satisfaction. Each scenario expressed a different level of some attribute of a purchase (good or bad, high or low, etc.) derived from the first study or from the literature as likely to impact customer evaluations. Each study is described in detail shortly (with theoretical reasoning and references provided). All scenarios are presented in Appendix C, with the phrases that differed across experimental conditions set apart in braces. The 10 survey pages were counter-balanced across participants, and all factors were between-subjects.

Each study was a two-factor design that yielded three omnibus hypothesis tests (two main effects and one interaction), for a total of $10 \times 3 = 30$ testable effects. We sought a "family-wise" alpha of .05, so we used a Bonferroni correction to obtain a per-test alpha of .05 divided by 30 = .00167. We did this
because the same participants were involved in all 10 experiments, yet the factors were not repeated measures, so there was no superior means of analyzing the data than that which we report. Note that in decreasing our alpha level from .05 to .00167, we are making it more difficult to obtain significant results (the tests are more conservative). Thus, we increase the likelihood that satisfaction and quality will not reflect empirical differences and will instead appear to represent the same construct. However, erring in this direction seems defensible, given scientific concerns for parsimony.

Two other features of this study warrant further mention. First, one of the factors in 9 of the 10 designs is whether the respondent was asked to make a judgment of quality or his or her anticipated satisfaction. That is, we manipulated which “consumer evaluation” construct each participant judged, with operationalizations of either “quality” or “satisfaction.” We are investigating the extent to which judgments are framed differently by the confrontation of the rating scale regarding quality versus the rating scale regarding satisfaction. This aspect of the design might appear somewhat novel, but logically it is no different from other more common measured factors (e.g., measurement order, etc.).

The second related feature of this study is that we chose to have different participants rate quality and satisfaction. The between- versus within-subjects design choice is delicate here. The obvious methodological shortcoming of a within-subjects design is that a participant encountering both rating scales might be induced by demand characteristics to distinguish their judgments when they might not have done so otherwise (e.g., “Why would they ask me to make both ratings unless they were different?”). Conversely, we were concerned that a within-subjects design would impede detection of differences between the two constructs due to method variance, participant heuristics, or laziness. Note that the between-subjects designs we chose involves different participants making the quality and satisfaction ratings. Like in any between-subjects design, our interpretive logic relies on random assignment and assumes that prior to experimental manipulations, the conditions are equated on extraneous factors. Nevertheless, because our goal is to examine differences, we wish to be sensitive to the possibility that any observed differences between quality and satisfaction may have stemmed from participant differences rather than reflecting differences between the constructs, and we return to the point after examining the results.

In interpreting the following results, recall that significant main effects for the purchase attributes indicate that our manipulations were effective. However, it is the interaction in each experiment that is of critical interest, because the interactions will support the argument that quality and satisfaction differ; that is, the impact of some purchase attribute on quality differs from that for satisfaction.

**Study 2.1.** In this first experiment, we examine the impact of disconfirmation (or gap). The literature defines quality as a function of the gap (cf.
Parasuraman et al., 1985) and satisfaction as a function of disconfirmation (cf. Churchill & Surprenant, 1982; Oliver, 1977, 1993; Tse & Wilton, 1988; Yi, 1990), both of which are relativistic judgments of experience and expectation. Thus, it is plausible that this interaction should not be significant; however, we wish to explore these predictions empirically.

Each participant evaluated one purchase scenario from a $3 \times 2$ factorial. The first factor was whether the described purchase experience exceeded, met, or fell short of the purported expectations said to be derived from word-of-mouth from friends (see Appendix C). The second factor was whether the respondent was asked to make a judgment of quality or satisfaction.

There was a significant main effect, $F(2, 35) = 52.58, p = .0001$, for disconfirmation. Participants' evaluations were poorest when the service provided fell short of their expectations, $M_{\text{fell short}} = 3.000$, and were significantly more favorable when the service either met or exceeded their expectations, $M_{\text{met}} = 5.429; M_{\text{exceeded}} = 5.750$; the latter two were not different: $F(1, 35) = .81, p = .3753$. There was no main effect for the "type of evaluation" made (i.e., quality or satisfaction), $F(1, 35) = .04, p = .8524$, and one would not be expected (such a main effect would imply that simply making ratings of quality or satisfaction yielded more favorable evaluations). The interaction was borderline significant, $F(2, 35) = 5.65, p = .0075$, recalling the alpha level of .00167, but as the plot in Figure 2A indicates, the interaction was primarily due to the difference between falling short versus meeting or exceeding expectations. There was also a difference in the impact of meeting expectations, compared to falling short of expectations. The asymmetry in the effect is that meeting expectations improved satisfaction more than quality (the difference between points $b$ and $d$ in Figure 2A, is greater than that between points $a$ and $c$: $F(1, 35) = 27.35, p = .0001$). Even though satisfaction and quality were not

![Figure 2A](image-url)

**FIGURE 2A** Selected Service Attributes $\times$ Customer Satisfaction and Service Quality interactions. Disconfirmation $\times$ Customer Satisfaction and Service Quality interaction, $F(2, 35) = 5.65, p = .0075$. Simple effects follow: Neither points $a$ and $b$, $F(1, 35) = 1.71, p = .1999$, nor $c$ and $d$, $F(1, 35) = 2.88, p = .0987$, nor $e$ and $f$, $F(1, 35) = 1.25, p = .2703$, are significantly different. Points $a$ and $c$ differ, $F(1, 35) = 15.07, p = .0004$, as do $b$ and $d$, $F(1, 35) = 63.73, p = .0001$. Points $c$ and $e$ do not differ, $F(1, 35) = .81, p = .3753$, and points $d$ and $f$, $F(1, 35) = .14, p = .7111$, do not differ.
judged differently when expectations were met (points \( c \) and \( d \) are not different), and they were not judged differently when expectations were not met (points \( a \) and \( b \) are not different), the slope from \( b \) to \( d \) is greater than that between \( a \) and \( c \); suggesting a greater impact on satisfaction than quality. However, the effect of exceeding expectations compared with meeting them did not impact quality differently from satisfaction.

Thus, there were no differences between quality and satisfaction with respect to disconfirmation in terms of falling short, meeting, and exceeding expectations. On the positive side, this finding supports both sets of researchers: those studying quality via the gap and those studying satisfaction via disconfirmation. As mentioned, however, meeting versus failing to meet expectations impacted ratings of satisfaction more than those of quality.

Study 2.2. The factors of reliability, consistency, and timeliness are thought to be important in the evaluation of services because the service encounter is said to be a process, with the service provision and the interaction between the service provider and client happening in real time (cf. Shostack, 1987). We examine these three attributes in the three experiments that follow.

In Study 2.2, we manipulate reliability. Following Parasuraman et al. (1989), we operationalized reliability as whether a promise given by a manufacturer or service provider is met during the provision of the service (see Appendix C). The second factor was whether respondents judged quality or satisfaction. There was a main effect for whether the service was reliable, \( F(1, 39) = 78.67, p = .0001 \). Ratings were more favorable when the promise was kept, \( M_{\text{promise kept}} = 5.850 \), than when it was not, \( M_{\text{not kept}} = 3.739 \). There was no main effect for whether the evaluation was of quality or satisfaction, \( F(1, 39) = 2.92, p = .0953 \), and there was no interaction between type of evaluation and reliability, \( F(1, 39) = 1.67, p = .2038 \). Thus, quality and satisfaction did not differ with regard to reliability.

These results resemble those of Study 2.1. Perhaps from the consumer perspective meeting a promise is similar to meeting an expectation, and quality and satisfaction appear to be essentially indistinguishable on these criteria. The finding on reliability is interesting because although it is one of the dimensions purported to be important to consumers' judgments of service quality (Parasuraman et al., 1985), it is not discussed explicitly by satisfaction researchers as a particularly important antecedent to satisfaction. These data suggest that reliability is important to either sort of evaluative judgment and therefore deserves greater recognition in customer satisfaction research.

Study 2.3. Reliability can also be thought of as consistency or repeatability (Shostack, 1987) and was operationalized as such for this study. The significant, \( F(1, 39) = 274.28, p = .0001 \), main effect indicated consistency was preferred to inconsistency, \( M_{\text{consistent}} = 6.000 \); \( M_{\text{not consistent}} = 2.087 \). There
was no main effect for the type of evaluation made (quality or satisfaction), $F(1, 39) = .30, p = .5873$, and there was no significant interaction, $F(1, 39) = 3.95, p = .0539$. Thus, consistent or inconsistent purchase transactions do not result in different quality versus satisfaction reactions.

**Study 2.4.** Another important dimension to Parasuraman et al. (1985) is timeliness—the quickness of response during the service encounter. The significant main effect, $F(1, 39) = 694.39, p = .0001$, indicated that quicker service was preferred to slower service, $M_{\text{quick}} = 6.286; M_{\text{not quick}} = 1.818$. There was no main effect for the type of evaluation made (quality or satisfaction), $F(1, 39) = 1.48, p = .2312$. We now also see an interaction between the service attribute variable (timeliness) and the type of evaluative judgment made by the respondents, $F(1, 39) = 17.55, p = .0002$. As the plot in Figure 2B indicates, timeliness affected perceptions of service quality, but it had an even greater impact on (i.e., is an even more important determinant of) customer satisfaction. Again, this finding is particularly interesting because timeliness is not a factor explicitly studied in the satisfaction literature, but these data speak in the consumer voice suggesting that it ought to be considered.

**Study 2.5.** The process aspect of services has additional implications beyond reliability, consistency, and timeliness. For instance, the purchase delivery process makes possible the customization of the purchase for a customer (cf. Surprenant & Solomon, 1987). Furthermore, interpersonal factors between the provider and customer become more important to the customer's evaluation of the service (cf. Crosby & Stephens, 1987; Iacobucci & Ostrom, 1993; Solomon, Surprenant, Czepiel, & Gutman, 1989). In this study, we examine customization (vs. standardization); in the next two, we examine the interpersonal factors of empathy and friendliness.

The significant main effect for customization, $F(1, 39) = 43.56, p = .0001$, indicated that respondents preferred customized service tailored to meet their

![FIGURE 2B](image-url)  
**FIGURE 2B** Timeliness × Customer Satisfaction and Service Quality interaction, $F(1, 39) = 17.55, p = .0022$. All four means are significantly different: Points $b$ and $c$ differ, $F(1, 39) = 503.53, p = .0001$, as do $a$ and $d$. $F(1, 39) = 228.69, p = .0001$. Points $a$ and $b$ differ, $F(1, 39) = 15.12, p = .0004$, as do $c$ and $d$. $F(1, 39) = 4.28, p = .0453$. 

special needs over standardized service provided to all customers, $M_{\text{customized}} = 5.435; M_{\text{standardized}} = 3.300$, but there was no main effect for the type of evaluation (quality or satisfaction), $F(1, 39) = .14, p = .7152$, and there was no interaction between these two factors, $F(1, 39) = .06, p = .8035$. Thus, customized service is perceived to be both more satisfying and of higher quality.

Study 2.6. The interpersonal factor examined in this study was empathy (cf. Parasuraman et al., 1985); a service provider's understanding of a customer's experience. A significant main effect, $F(1, 37) = 18.11, p = .0001$, indicated that empathic service providers were preferred to those less understanding, $M_{\text{empathic}} = 5.947; M_{\text{no empathy}} = 2.636$, but there was no main effect for the type of evaluation made, $F(1, 37) = 1.05, p = .3117$, and there was no interaction, $F(1, 37) = .11, p = .7453$.

Study 2.7. The second interpersonal factor was friendliness, and the results were similar to those for empathy. The significant main effect, $F(1, 38) = 363.84, p = .0001$, indicated friendly service personnel were preferred to less friendly service providers, $M_{\text{friendly}} = 5.842; M_{\text{not friendly}} = 1.435$, but both the main effect for the type of evaluation made, $F(1, 38) = .00, p = .9819$, and the interaction between these factors, $F(1, 38) = 1.73, p = .1957$, were not significant. The similarity between these results and those for empathy suggest that these factors were perhaps only slightly different operationalizations of interpersonal effects.

Study 2.8. Price, one of the four fundamental mix variables, is frequently studied in its own right (e.g., Kalwani & Yim, 1992). It is also important to customers' judgments of satisfaction and quality, because price and the goodness of the purchase (e.g., quality) are thought to be combined into evaluations of value (see Zeithaml, 1988).

In this study, there was a borderline main effect for price, $F(1, 38) = 6.86, p = .0126$, indicating that high-priced services were judged somewhat more favorably than less expensive services, $M_{\text{high price}} = 5.000; M_{\text{low price}} = 4.087$. There was no main effect for the type of evaluation made (quality or satisfaction), $F(1, 38) = 2.29, p = .1385$. However, there was a significant interaction, $F(1, 38) = 24.06, p = .0001$, as the plot in Figure 2C indicates, price had no impact on satisfaction, but affected judgments of quality. Specifically, ratings were most favorable for higher priced purchases than less expensive services, so we might conclude that price serves as a cue to judgments of service quality. Thusfar, this is the second factor that has yielded different effects on quality and satisfaction.

Study 2.9. Many services marketers find a dramaturgy analogy appealing, referring to the aspects of service provision visible to a customer as the front
FIGURE 2C Price × Customer Satisfaction and Service Quality interaction, $F(1, 38) = 24.06, p = .0001$. Points $b$ and $c$ are not different, $F(1, 38) = 2.64, p = .1126$, but points $a$ and $d$ are, $F(1, 38) = 28.05, p = .0001$. Points $a$ and $b$ are different, $F(1, 38) = 18.80, p = .0001$, as are $c$ and $d$, $F(1, 38) = 6.36, p = .0160$.

FIGURE 2D Backstage × Customer Satisfaction and Service Quality interaction, $F(1, 38) = 8.01, p = .0047$. Points $a$ and $d$ do not differ, $F(1, 38) = 2.37, p = .1318$, but $b$ and $c$ do, $F(1, 38) = 6.05, p = .0186$. Points $a$ and $b$ differ, $F(1, 38) = 5.56, p = .0236$, but $c$ and $d$ do not, $F(1, 38) = 2.77, p = .1044$.

...stage and those supportive aspects occurring behind the scenes, out of the customer's line of vision, as the back stage (cf. Grove & Fisk, 1983; Lovelock, 1991). Most of the aspects we have studied thus far might be classified as front stage (e.g., perceptible friendliness of service providers). In this study, we focus on the back stage; operationalizing an aspect invisible to customers as the effectiveness of accounting software that a hotel staff uses.

The main effects were not significant: for back stage, $F(1, 38) = .44, p = .5116$; for type of evaluation, $F(1, 38) = .18, p = .6757$. However, the interaction between these factors was borderline significant, $F(1, 38) = 8.01, p = .0074$. As the plot in Figure 2D indicates, improvements to a back stage feature of a service system had no impact on respondents' judgments of satisfaction, but being informed of a poor back stage feature (in effect making it become front stage) resulted in more negative perceptions of service quality.

Study 2.10. In the previous nine experiments, we manipulated an antecedent service attribute and examined its possibly different impact on quality and satisfaction. In this final experiment, we study the impact of quality and satisfaction on the consequence of a customer's purchase intentions. We instructed the participants to consider they had purchased a high-quality service or (to other participants) one that induced customer satisfaction, implicitly letting the participants define these terms themselves. We then asked them to estimate how likely they would be to return for repurchase. Given that the literature on both quality and satisfaction purport that purchase intentions are a consequence of evaluations (cf. Churchill & Surprenant, 1982; Oliver, 1977; Parasuraman et al., 1985; Tse & Wilton, 1988; Yi, 1990), we would predict this
factor, likelihood to purchase, should not yield different results for high-quality or highly satisfactory purchases.

The significant main effect for the goodness of the service, $F(1, 39) = 182.32, p = .0001$, indicated that good evaluations led to greater likelihoods of repurchase, $M_{\text{good service}} = 6.400; M_{\text{bad service}} = 1.826$, but there was no significant main effect for the type of evaluation given (quality or satisfaction), $F(1, 39) = .004, p = .9521$, and there was no interaction between these factors, $F(1, 39) = .28, p = .5988$. Thus, either high quality or high satisfaction appears equally likely to lead to intentions of purchase.

Summary

From these 10 experiments we found two service attributes to distinguish judgments of quality and satisfaction: timeliness and price. The former had a greater impact on satisfaction, the latter on judgments of quality. A third attribute deserves further study, given that the interaction was of borderline significance. That finding indicated that although knowledge of the back stage did not have an impact on satisfaction, it did affect perceptions of quality.

In addition, we found that, overall, the disconfirmation-gap concept is a plausible antecedent for both quality and satisfaction and that purchase intentions are a plausible consequence for either; both findings are consistent with these literatures, as depicted in Figure 1. Note that neither distinguishes quality from satisfaction. Finally, we also found that certain service attributes were preferred to others (meeting or exceeding expectations, promises being kept, consistency in service orientation, timeliness, customization of services provided, and empathy and friendliness of service provider). These differences also indicated that we were successful in manipulating the service attributes.

WITHIN- Versus BETWEEN-SUBJECTS DESIGNS

Lastly, we wish to return to the concern that the between-subjects designs may enhance apparent (or otherwise nonexistent) differences between quality and satisfaction. This possibility is interesting, given that very few differences were actually found; it would suggest that quality and satisfaction should probably be considered as a single construct (an outcome with no more or less value than considering them as distinct). To explore this issue further, we ran a within-subjects version of these studies on a new sample of 77 participants (drawn from the same population). We offer the results and a caution.

Regarding the results, none of the within-subjects versions of these studies yielded significant interactions (whereas there had been three in the between-subjects version). One factor that yielded an interaction in the between-subjects version (price) was borderline in the within-subjects treatment, $F(1, 73)$
\[ 3.45, p < .06, \] with means that were in the same direction—price had a greater impact on quality than on satisfaction. As such, these between-subjects results would be positioned as a liberal demonstration of the dimensions on which quality and satisfaction differ. (The within-subjects approach cannot be said to be conservative because it has its own issues, which we discuss momentarily.) Given the scientific penchant for conservatism, the best conclusion might be that quality and satisfaction are a single construct.

Regarding the caution, we note that the difficulty we anticipated regarding the within-subjects design arose. The ratings of quality and satisfaction in this study (where each participant made both judgments) were very highly correlated. The mean correlation between satisfaction and quality across the studies was \[ r = .909. \] It is not surprising that respondents did not make much distinction between two ratings that apparently they thought of as synonymous, but it highlights the delicacy of methodology required if researchers wish to distinguish the two consumer evaluations.

**SUMMARY AND DISCUSSION**

Quality and satisfaction have been conceptualized similarly in the literature and therefore might be parsimoniously considered as one construct. However, some researchers conceptualize differences, and it might be more interesting theoretically to consider the two as distinct because this point of view would necessitate hypothesizing distinct positions in presumably more complex nomological networks.

In this article, we pursued the consumer voice; we presented two studies in which we allowed respondents to decide whether they thought the terms quality and satisfaction were similar or different. One study was a qualitative analysis of open-ended survey responses, the other, a set of experiments. Given the two different methods, it was encouraging that there was some convergence across studies; for example, no differences between quality and satisfaction were found in either study for disconfirmation, keeping promises, customization, empathy, friendliness, or purchase intentions. Moreover, the results provide support for both the quality and satisfaction literatures in their positing a relative judgment of experiences versus expectations as an antecedent and purchase intentions as a consequence. However neither finding distinguishes quality from satisfaction. Although null results remain open to alternative explanations, convergent lack of differences resulting from multiple studies with complementary methodologies are less so.

Turning to the differences, both types of data supported several distinctions between quality and satisfaction. The nomological network in Figure 3 is essentially extracted from the much larger number of purchase attributes tested in this article and posited in the literature. These differences are empiri-
### Antecedents

- Price
- Back-Stage
- Expertise

### Focal Constructs

- Service Quality
- Customer Satisfaction
- Timeliness
- Service Recovery
- Physical Environment

**FIGURE 3** Conceptualizing service quality and customer satisfaction: Uniqueness among antecedents.

Cal distinctions (of kind or amount) between quality and satisfaction from the customer's point of view. As per the data, we indicate the purchase attributes of price, back stage, and expertise as causal antecedents more likely to affect judgments of service quality, and timeliness, service recovery, and physical environment as those more likely to affect perceptions of customer satisfaction. Specifically, differences in kind were as follows: price and back stage affected quality but not satisfaction. Differences in amount were as follows: timeliness, service recovery, and physical environment affected satisfaction more than quality, whereas expertise affected quality more than satisfaction.

It would be interesting to entertain possible theoretical generalizations regarding what price, back stage, and expertise may have in common that made them more descriptive of quality, whereas timeliness, service recovery, and physical environment were factors more diagnostic of satisfaction. Perhaps the quality factors are those primarily under the control of management—managers set prices, design and implement the flowchart of service delivery including the behind-the-scenes systems, and are responsible for the expertise of their
service providers, which motivates policies regarding employee selection and training. In contrast, perhaps the satisfaction factors are those that impact the experiential process aspects of the service purchase from the consumer's perspective: The customer's on-site service experience is impacted by the speed of the service, how well recovery takes place if an error is made, and the appearance of the environment in which the encounter occurs. The generalization is not perfect, given that management presumably has control over aspects of the physical environment, but we are suggesting that the environment has greater importance on the customer experience.

Thus, perhaps the difference between quality and satisfaction mirrors managerial versus customer concerns; a manager and service-providing firm tries to provide "high-quality" service, and a customer experiences the service encounter and is "satisfied" or not. Characterizing quality as within the domain of managers and satisfaction as the evaluative reaction of customers would offer a clear basis for distinction: "Quality" programs involving total quality management, 6σ (e.g., Burr, 1976), and so forth would focus on improving the managerially controllable aspects of the service-delivery system, and measures of customer satisfaction would capture the consumer reaction. Presumably, if quality programs were initiated based on marketing research—that is, the changes were market driven and customer oriented—then quality improvements should lead to customer satisfaction. Perhaps it is when management is least in tune with customers and is doing the poorest job at "marketing" that quality differs from satisfaction: A firm may provide "high-quality" service that nevertheless does not "satisfy the customer" because the properties improved on do not matter to their consumers.

Finally, we might also note that neither duration (short-term vs. long-term) nor affect-cognition, two concepts recently proposed to distinguish quality and satisfaction, were mentioned by respondents in their differentiations of quality and satisfaction judgments in our studies. As we had stated earlier, this does not negate the possible importance of such theoretical differences for future conceptual work. However, these distinctions were not reflected in the consumer perspective, and to the extent that research on quality or satisfaction should reflect the marketplace of those whose opinions are being conceptualized, we believe the consumer voice is important.

ACKNOWLEDGMENTS

This study was supported by National Science Foundation Grant SES-9023445.

We are grateful to Bridgette Braig, Dipankar Chakravarti, Brian Sternthal, Terri Swartz, and marketing seminar participants at Ohio State University for their helpful comments about this research. We thank Jennifer E.
REFERENCES


APPENDIX A

Stimuli for Study 1

The following questions were the stimulus for Study 1:

"Think about when, as a customer, you have recently had an encounter with an employee of a service firm (e.g., airline, hotel, restaurant, etc.) that you would characterize as a \{"high quality"/"low quality\} service, yet you were \{dissatisfied/satisfied\} as a customer anyway. Please describe this service encounter.

a) When did the incident happen? With which service firm?
b) What specific circumstances led up to the situation?
c) Exactly what did the service provider say or do?
d) What about the encounter made you believe it to be of \{high/lowl quality\}?
e) What about the encounter made you feel it was \{dissatisfying/satisfying\}?"

These stimuli were a modification from Bitner et al. (1990), who had asked, "Think about when, as a customer you had a particularly satisfying (dissatisfying) interaction with an employee of an airline, hotel, or restaurant." Our questions (a) through (c) replicate those asked by Bitner et al., and where they asked question (e), we included both (d) and (e).

APPENDIX B

Brief Version of Coding Scheme Used in Study 1

Defn = definition; Ex = example; Ref's = references.
A. Title: Empathic, emotional or intangible response to a mistake or problem

Defn: When a mistake occurs and the respondent focuses on the emotional response offered or not offered, showing empathy or remorse, understanding customer frustration.

Ex: "How they seemed to feel about the mistake."

B. Title: Tangible response to a mistake or problem or lack thereof
De'n: When a mistake occurs and the respondent focuses on tangible compensation or free merchandise offered or not offered.
Ex: "Gave us a free dessert."

C. Title: Customized requests or preferences: responding to, not responding to, or mentioning them
De'n: Respondent focuses on a specific or personal need.
Ex: "I asked for flowers in the room, and that was no problem."
Refs: Surprenant and Solomon (1987)

D. Title: Recognizing (or not recognizing) customer or customer value
De'n: Customers like to be recognized immediately, listened to, and shown courtesy and politeness. (Note this is not friendliness.)
Ex: "They clearly didn't value my business."
Refs: Lovelock (1991)

E. Title: Speed/response or lack thereof
De'n: Focus on how quickly or slowly the service or response was.
Ex: "It took her way too long to answer my question."
Refs: Parasuraman et al. (1985)

F. Title: Ease of use or lack thereof
De'n: Respondent focuses on how easy, simple, convenient the product/service is.
Ex: "It was hard to do, but we had help."

G. Title: Friendliness or niceness or rudeness, "touchy-feely"
De'n: Respondent focuses on niceness or charm of the service provider.
Ex: "The whole staff was about as nice as you could get."

H. Title: Did/did not do job well—namely, expertise or professionalism or competence
De'n: Respondent talks generally about the ability of the provider or company to do the job well. (Use this only if no other more specific categories apply.)
Ex: "The mechanic really seemed to know about foreign cars."
Refs: Crosby and Stephens (1987), Parasuraman et al. (1985)

I. Title: Price/value or lack thereof
De'n: Any focus by the respondent on cost, payment, or money.
Ex: "Price was low."

J. Title: Information or lack thereof
De'n: Respondent focuses on the amount of information provided.
Ex: "The salesman couldn't tell me some of the important aspects of the product."
Refs: Lovelock (1991)

K. Title: Good/bad physical aspects of the environment, provider, or product
De'n: This includes any mention of the surroundings or ambiance (e.g., decor, food, dress, and safety).
Ex: "The dining room was elegant."
Refs: Bitner (1992)

L. Title: Promises, lying, deception
De'n: Respondent focuses on a promise that was made, kept or not kept, or focuses on the service provider's deception.
Ex: "The pilot said we would be only 20 minutes late."
Refs: Parasuraman et al. (1985)
M. Title: Reputation/recommendations
   Defn: Respondent focuses on the reputation of the company or the provider or on things that others have said about the company. Sometimes, expectations based on reputation are also mentioned.
   Ex: "I heard from my neighbor that it was a good place."
   Ref’s: Lovelock (1991)

N. Title: Comparison with previous experience with this provider
   Defn: Respondent mentions previous experience with this company or provider and makes a comparison or mentions an expectation.
   Ex: "Not what we thought would happen, given our previous visit to the hotel."

O. Title: Mention of future encounters with provider
   Defn: Respondent mentions intentions regarding future encounters with this provider.
   Ex: "Will definitely come back."

P. Title: Customer focuses on own affect
   Defn: When customer’s key focus is on how he or she or fellow customers felt during the encounter.
   Ex: "It made me blow my top."
   Ref’s: Mano and Oliver (1993), Oliver (1993), Westbrook and Oliver (1991)

Q. Title: Customer focuses on own behavior
   Defn: When customer’s key focus is on how she or he or fellow customers acted during the encounter.
   Ex: "That’s when we just left."

R. Title: Neutral description of encounter or story or Too general or cryptic to categorize
   Defn: Use this when the sentence is clearly part of a story, but it is difficult to decide how it relates to the encounter.
   Ex: "Before we knew it, the glass was empty."

APPENDIX C

Stimuli for Study 2

Study 2.1: Disconfirmation. (References in body of article.) Imagine you have just discussed purchasing an insurance policy with an insurance agent. You had heard about the insurance agent from some friends before you went, so you had certain expectations about how the meeting would go. Your visit {met/exceeded/fell short of} your expectations. How would you rate the quality of the insurance agent?

very low quality  1 2 3 4 5 6 7  very high quality
-OR- very dissatisfied  1 2 3 4 5 6 7  very satisfied

Study 2.2: Reliability (i.e., a promise is kept). The hotel at which you’ve booked reservations for a vacation has promised that you will never have to wait more than a minute in line to check in. You arrive at the hotel and find that the promise {is/is not} kept. You {do not/do}
have to wait in line for more than a minute to get your room key. Given this experience, how would you rate {the quality of/your satisfaction with} this hotel?

**Study 2.3: Consistency.** The insurance agent who holds your car insurance policy seems to have {consistent/inconsistent} service. You find that the receptionists, the agents, and the claims adjusters all have {the same level/quite different levels} of customer orientation. In addition, the insurance agency seems to have been very {consistent/inconsistent} in its efforts to make it as easy as possible to make claims and yearly payments. Given this experience, how would you rate {the quality of/your satisfaction with} this insurance agency?

**Study 2.4: Timeliness, responsiveness.** Imagine that you've had to put in a small claim on your automobile insurance. Processing the claim took {less/more} time than you thought it was going to take; the insurance agent {seemed to take care of the claim quickly/seemed to be slow in taking care of the claim} and was very {responsive/unresponsive} to your concerns. How would you rate {the quality of/your satisfaction with} this service?

**Study 2.5: Customization.** A new insurance agency is undertaking measures to {customize/standardize} its offerings. It {can provide a wide variety of specialized policies to meet the individual needs of each of its clients/provides only a few set policies that clients can choose}. How {would you characterize the quality of this service/satisfied would you be with this service}?

**Study 2.6: Empathy.** While it is parked in a supermarket parking lot, your car is dented badly by a hit-and-run driver. You call your insurance company immediately, and find that the agent {is/is not} very understanding. She {seems to/doesn't seem to} understand how frustrating it is to have a car accident, especially one that isn't your fault. Given this experience, how would you rate {the quality of/your satisfaction with} this insurance agency?

**Study 2.7: Friendliness.** While staying at a hotel, the desk clerk was {very pleasant/rude} and the concierge was {friendly and courteous/unfriendly and discourteous} while recommending some things for you to do and see while in town. How {would you rate the quality of/satisfied would you be with} this hotel?

**Study 2.8: Price.** Imagine that you are planning to take a vacation and need to make reservations at a hotel. There are several that are conveniently located, however, one is priced 20–25% {more/less} than the others. If you were to choose this hotel, {what level of quality would you expect it to be/how satisfied would you be}?

**Study 2.9: Back-stage.** The hotel at which you've made vacation reservations {has updated/has decided not to update} its accounting software {in order to make/that would have made} weekly reports more organized and succinct. The transition to the new software {was difficult/would have been too difficult}, {but the finance department decided it was necessary/so the finance department decided it was unnecessary}. How would you rate {the quality of/your satisfaction with} this hotel?

**Study 2.10: Likelihood of purchase (as a consequence).** Imagine you have just spent a couple of nights at a hotel. You {were a "highly satisfied (dissatisfied) customer"/thought the visit was “very high (low) service quality”} and you filled out the in-room questionnaire
accordingly. Think about \{your level of satisfaction/this level of service quality\}; if you were to return to this city and needed a hotel again, how likely is it you would stay at this hotel again?

not very likely at all 1 2 3 4 5 6 7 very likely

Accepted by Dipankar Chakravarti