DUNCAN I. SIMESTER, JOHN R. HAUSER, BIRGER WERNERFELT, and ROLAND T. RUST*

A multinational firm uses sophisticated, state-of-the-art methods to design and implement customer satisfaction improvement programs in the United States and Spain. Quasi-experimental analysis reveals a complex and surprising picture that highlights (1) implementation issues (empowerment), (2) a construct of residual satisfaction not captured by customer needs, and (3) the managerial need for combining nonequivalent controls and nonequivalent dependent variables.

Implementing Quality Improvement Programs Designed to Enhance Customer Satisfaction: Quasi-Experiments in the United States and Spain

Although branches of the marketing literature vary in their analyses and definitions of constructs, they generally agree that if a firm were to improve its products and services to fulfill important customer needs, that firm would enhance its customers' satisfaction and its future profits (Anderson and Sullivan 1993; Fornell 1992; Hauser, Simester, and Wernerfelt 1994, 1996, 1997; Rust, Zahorik, and Keiningham 1995; Zeithaml, Parasuraman, and Berry 1990). Such prescriptions are common in textbooks and monographs and are supported by several laboratory studies and many compelling anecdotes of product and service quality successes. However, there are few published field experiments (or quasi-experiments) designed to test whether firms can implement quality improvement interventions that lead to measurable improvements in customer satisfaction. One notable exception is Bolton and Drew's (1991) description of GTE's attempt to improve telephone service.

Widespread acceptance of the relationship is evident in the growing popular literature on quality and in the reliance on customer satisfaction measures in new product development and employee compensation (Anderson, Fornell, and Lehman 1994). In contrast, after initially accepting the wisdom of implementing new quality interventions, firms are now beginning to demand explicit justification for their investments. Consulting firms and industry commentators are encouraging this trend with a series of studies questioning the benefits of quality interventions (Howe, Gaeddert, and Howe 1995; Rust, Zahorik, and Keiningham 1995):

> A study by the American Quality Foundation and Ernst & Young suggests that many companies are wasting their efforts in trying to improve quality. The consulting firms of AT Kerney and Arthur D. Little present equally disappointing findings in two separate studies: (1) 80% of more than 100 British firms reported "no significant impact as a result of TQM [total quality measurement]" and (2) almost two-thirds of 500 U.S. companies saw "zero competitive gains." (Anderson, Fornell, and Lehman 1994, p. 53)

A scientific evaluation of an actual quality improvement program provides an important contribution to this debate. In this article we describe two quasi-experiments undertaken by a technically sophisticated, $2-billion international firm seeking to evaluate a new quality and customer satisfaction intervention. For this article we disguise the firm by calling it KemTek. The program was initiated by a CEO-led task force that had determined that enhancing customer sat-

*Duncan I. Simester is Associate Professor of Management Science (e-mail: simester@mit.edu), John R. Hauser is Kirin Professor of Marketing (e-mail: jhauser@sloan.mit.edu), and Birger Wernerfelt is Professor of Management Science (e-mail: bwerner@mit.edu), all at the Sloan School of Management, Massachusetts Institute of Technology. Roland T. Rust is Madison S. Wigginton Professor of Management and Director of the Center for Service Marketing, Owen Graduate School of Management, Vanderbilt University (e-mail: roland.rust@owen.vanderbilt.edu). This research was funded by the International Center for Research on the Management of Technology and the Center for Innovation in Product Development. The authors thank Ning Peng, Lisa Tener, Christine Forte, and Robert Klein, who participated in the design and implementation of the quality improvement effort and the quasi-experiment. The authors also acknowledge the generous contributions of time and data by the firm that is the subject of this study. Applied Marketing Science, Inc., in Waltham, Mass., conducted the qualitative research. To interact with colleagues on specific articles in this issue, see "Feedback" on the *JMR* Web site at www.ama.org/pubs/jmr.

isfaction was key to the firm's survival. The task force commissioned state-of-the-art marketing research to determine important customer needs and adopted widely used quality tools to design a customer satisfaction improvement program to make its products and services meet these important customer needs. It hoped the improvements would enhance customer satisfaction and therefore the firm's long-term profitability in its business-to-business market.

The implementation is unique from the perspective of field research because this firm, with its science-oriented culture, sought to measure the impact of the program by (1) designing parallel quasi-experiments in two countries and (2) investing approximately two years and $500,000 in data collection. Given the significant capital investment required to implement similar programs in all its divisions, the firm sought to determine whether this investment was justified. By describing the development and implementation of the firm's intervention and by providing data on the outcome of the quasi-experiments, we seek to add insight to the relevant scientific literature in marketing and to provide a window on the implementation of one large-scale, market-driven customer satisfaction intervention in industry.

The study confirms the basic premise: KemTek's efforts led to measurable and enduring improvements in satisfaction of the targeted needs. Beyond this, however, the data contain a number of surprises. At least three commonly held beliefs are questioned. First, although we (and KemTek) believed ex ante that the treatments in the two countries were equivalent, this appears not to have been the case. Seemingly minor and inconsequential differences in the manner in which employees were empowered appear to have been quite important. Second, although we used state-of-the-art methods to elicit customer attributes (and evidence suggests that the list of attributes was fairly exhaustive), some of the improvement in overall satisfaction was due to factors not captured in attribute satisfaction. This suggests the need for academic research to improve the measurement (and understanding) of the determinants of overall satisfaction. Third, despite relatively large sample sizes, the nonequivalent treatment and control groups were not sufficient to identify significant effects in both quasi-experiments. Fortunately, we had data available to incorporate a nonequivalent dependent variables design with the standard pretest–posttest untreated control group design to enhance the power of the analysis (Cook and Campbell 1979, pp. 249, 261). Together these three lessons may help explain why industry commentators and firms themselves are starting to question the wisdom of quality interventions. If these three effects are not taken into account, industry might be led to erroneous conclusions that some quality and customer satisfaction interventions do not work when a more complete analysis would suggest otherwise.

QUASI-EXPERIMENTAL DESIGN AND MEASURES

KemTek implemented a quasi-experimental design that included an extensive array of controls, including use of (1) pretest and posttest measures from the same panel of customers, (2) nonequivalent control groups, (3) nonequivalent dependent variables, and (4) replication in separate countries (chosen by KemTek to vary in the amount of competition). Using Cook and Campbell's (1979) notation, we depict the design in each country as shown in Figure 1.

Figure 1
QUASI-EXPERIMENTAL DESIGN

O_{1A}	$\times$	O_{2A}
O_{1B}		O_{2B}
O_{1A}		O_{2A}
O_{1B}		O_{2B}

Measurement activities are denoted by O, X denotes treatment activities, and the dashed line distinguishes between treatment and control regions. That is, KemTek chose a set of treatment regions in each country (above the dashed line) and implemented the customer satisfaction improvement programs to every customer that it could reach in the treatment regions. KemTek did not implement the program or undertake any special activities to improve customer satisfaction in the control regions (below the dashed line). Identical measurement instruments (O) were fielded before (O_1) and after (O_2) the intervention. The measurement instruments included two sets of dependent variable scales, one that was expected to change because of the treatment (O_A) and one that was not expected to change (O_B). (For a detailed discussion of the use of two related but different sets of dependent variables, see Cook and Campbell 1979, p. 261.) The design was replicated in two countries. Because the interventions were not identical in both countries, we consider the implementations as two separate but related quasi-experiments. Technically, this is equivalent to allowing a country variable to interact with every other variable and variable interaction.

The implementation of the quasi-experiments is detailed in a technical appendix, which is available from the authors. The implementation included the following steps: (1) product and market selection, (2) identification of customers' needs, (3) design of pretest and posttest measurements, (4) design of the customer satisfaction improvement program (experimental intervention) to improve customer satisfaction by targeting key customer needs, and (5) implementation of the customer satisfaction improvement program. With the exception of the posttest measures (which occurred after the implementation of the improvement program), this list represents an approximate chronological ordering of the activities.

KemTek sells a variety of integrated systems based on chemical, electronics, and materials technology. Some of its products are sold in the business-to-business market, some directly to end consumers, and some to intermediate customers (retailer/manufacturers) that use KemTek's product as raw material to produce finished goods for the end consumer. At the time we became involved, a CEO-led task force had decided that KemTek's profits were stagnating and that the firm could increase long-term profits by undertaking programs to increase customer satisfaction.[1] This

[1]By customer satisfaction, the task force referred to a long-term customer attitude that would enable KemTek to retain customers profitably. This definition differs from that used by, say, Bolton and Drew (1991), who refer to customer satisfaction as a transitory judgment based on a single transaction. Cronin and Taylor (1994, p. 126) refer to the long-term attitude as "service quality." KemTek's definition is similar to cumulative satisfaction, which is a proxy for future economic returns (Anderson, Fornell, and Lehmann 1994, p. 54). For the remainder of the article, we adopt KemTek's definition but caution the reader that the literature varies in its use of the words "customer satisfaction" (Yi 1990).

would be a significant capital investment over five years, so the task force decided to test the intervention using a major product/market chosen from KemTek's line of businesses.

The market that best matched the criteria was a market in which KemTek's products were used by small retail stores to produce a finished product for the end consumers. Although we cannot name the product category, the business-to-business customers (retailers) in this category would be analogous to tailors who produce finished apparel from cloth and sewing machines. In KemTek's market, the final item was sold for approximately $10, whereas the intermediate product cost the retailer approximately $1.50 per item. Production equipment (analogous to sewing machines) was a one-time capital cost for the retailer. The quality of the final item depended on the manner in which the intermediate product was used and stored by the retailer. Those retailers with greater expertise in these tasks produced higher-quality finished goods. KemTek's task force believed that training in the use and storage of the intermediate product combined with the maintenance and use of production equipment would greatly improve the quality of the final item and increase customer satisfaction. (One of the authors underwent training on KemTek's products and can attest both that producing high-quality output requires expertise and that training improves the judged quality of the output.) Here customer satisfaction refers to the satisfaction of the retailers, which can now produce a higher-quality product and, hopefully, increase their profitability.

KemTek's activities were global, so the task force interviewed management at KemTek's corporate office to select two countries in which to implement the intervention. The United States and Spain were selected. There were many similarities between the United States and Spain. KemTek sold identical products in both countries, the product was used for the same purposes, it was distributed through independent distributors, it was used to produce the same final product, and this final product was produced and sold by similar types of retailers. KemTek's management structures were also similar in the United States and Spain—the countries shared the same senior management, and local managers were company employees in regular contact with the corporate office.

Although KemTek's products were branded, retailers did not purchase directly from KemTek. They purchased from independent, nonexclusive distributors that supplied retailers with a variety of products from both KemTek and other firms. The distributors neither dealt exclusively with KemTek nor enjoyed exclusive markets. The distributors employed their own sales staff and, typically, did not offer training in the use or storage of KemTek's products. KemTek's own direct sales force did not have frequent contact with the retailers, concentrating instead on larger-volume customers in other product categories. (Before the intervention, the Spanish sales staff may have visited their customers relatively more than the U.S. sales staff. However, such visits were not made frequently.) The task force believed that a carefully designed program to call on retailers represented an opportunity to enhance retailers' satisfaction.

Notwithstanding these similarities, the two countries differ. One important difference is competition. At the time of the intervention, KemTek enjoyed an effective monopoly in the United States. However, in other countries, a global competitor was beginning to offer products that competed with KemTek's products. KemTek saw the entry of this competitor as a major threat to its future profitability and believed that defending KemTek's worldwide markets was a primary strategic objective for the firm. To gather data for strategic decisions, the firm chose to run one quasi-experiment in a country in which there was no competition (the United States) and another quasi-experiment in a country in which the competitor had already entered (Spain). If the findings were similar, they would represent evidence that the effect of the intervention was relatively insensitive to changes in the intensity of competition. If the findings were different, the competitive environment might help explain the observed disparities.

In Spain, the competitor had entered the market approximately two years before the intervention and had begun to offer a substitute product at a slightly lower price. Some Spanish retailers were not yet aware of the availability of the competitive brand, but all were aware of KemTek's brand. KemTek believed that supply constraints had restricted the competitor's ability to enter the domestic U.S. market. It correctly predicted that entry would not occur before completion of the study. On the basis of its long experience in these two markets, KemTek believed that the presence of competition would be a much larger factor than would other differences between the United States and Spain.

Three U.S. cities, New York, San Francisco, and Dallas, were assigned to treatment status, whereas Chicago, Los Angeles, and Miami were assigned to control status. In Spain, Barcelona and Malaga were designated treatment regions, and Madrid and Alicante were designated control regions. KemTek anticipated that these divisions would yield large samples of approximately equivalent customers.

The two-step voice-of-the-customer analysis described by Griffin and Hauser (1993) was used to identify important retailer needs. This process yielded a list of 17 needs. KemTek invested significant effort in identifying these needs. On the basis of its experience in the market and the voice-of-the-customer theory, KemTek believed that a combination of the 17 needs would almost completely explain the systematic variance in overall satisfaction. The proprietary nature of the data prevents us from publishing a complete description of each need; however, we provide an edited description subsequently.

Pretest and posttest measures were collected through telephone interviews. The pretest measures were collected approximately six months before the intervention and the posttest measures were collected approximately six months after the intervention. The same retailers answered both the pretest and the posttest questions. The interviews included the following groups of questions: (1) brand awareness, (2) overall satisfaction, (3) satisfaction with each of the 17 (retailer) customers' needs, (4) perceived importance of each of the 17 customers' needs, and (5) demographics and store description. A total of 327 and 224 complete responses were collected in Spain and the United States, respectively, distributed across the test and control cities as shown in Table 1.

The house-of-quality methodology was used to guide the design of the intervention programs. On the basis of the importance measures, the cost and feasibility of actions that KemTek might take, the estimated impact of the actions, and

Table 1
DISTRIBUTION OF RESPONSES

Treatment Regions		*Control Regions*	
New York	55	Chicago	89
San Francisco	36	Los Angeles	21
Dallas	8	Miami	15
Total	99	Total	125
Barcelona	64	Madrid	108
Malaga	75	Alicante	80
Total	139	Total	188

discussions with the local managers in both the United States and Spain, an interfunctional team decided to focus on 5 of the 17 retailer needs. The remaining 12 needs can be categorized further. Five are entirely distinct from the targeted needs and were unlikely to be affected by the intervention. In particular, the intervention programs did not alter the price of the product; change the durability, look, or portability of the equipment; or provide any advertising support to the retailers. The remaining 7 needs are less distinct, so that the intervention could have had an ancillary effect on these needs. For ease of exposition we use the terms "targeted," "ancillary," and "distinct" to categorize the different needs. We summarize these categories in Table 2.

The difference in the anticipated impact of the intervention on the 17 needs provides the nonequivalent dependent variables control (Cook and Campbell 1979). Use of the nonequivalent dependent variables enables us to control for any unobserved change in overall satisfaction with KemTek, perhaps due to competition in Spain, that increases or decreases satisfaction with all needs—targeted, ancillary, and distinct. Such changes in perceptions of all needs are often called "halo effects" (Beckwith and Lehmann 1975, 1976; Crosby and Stephens 1987). Because the distinct needs were not targeted by the intervention and were not expected to change differentially in the treatment or control regions, we can use them to control for both unobserved halo effects and other unobserved ecological changes.

Decisions on the final design of the intervention were made by committees comprising each country's local management, the authors, the task force, and senior management from KemTek's corporate office. These committees judged that the intervention would improve perceptions with respect to the targeted needs and that the net effect would increase long-term profits. The specific intervention was a training program designed to help retailers (1) improve their use and storage of KemTek's intermediate product and (2) set up and maintain their production equipment to make the best use of KemTek's product. Training procedures and collateral materials (procedure checklists, product samples, replacement parts, and accessories) were each designed to focus on one or more of the five targeted needs. The team intended that the effect of the intervention would be the same in both the United States and Spain but that the details would be optimized to the local situations in each country.

IMPLEMENTATION OF THE CUSTOMER SATISFACTION IMPROVEMENT PROGRAMS

United States

Eleven management volunteers, whose prior experience provided them with the necessary expertise, implemented the experimental treatment. This effort represented incremental resources invested by KemTek (the existing sales force continued its normal activities).

To standardize the treatment, the implementation team participated in a group training session and received detailed scripts and procedure manuals. Implementation of the program began four months after completion of the pretest measures and started with a telemarketing call by each representative to an assigned sample of approximately 20 retailers. The goal of this first contact was to establish an initial relationship with the retailer and schedule a convenient time for a site visit. The representatives were instructed to assure the retailers that the purpose of the visit was neither to collect data nor to induce a purchase but rather to offer assistance in the use of KemTek's product.

Actual field visits of approximately one hour were made to 179 of the retailers in the treatment regions that participated in the pretest measures. The field visits began with the representative's determining the quality of the manufactured item at the start of the visit. The representatives then described storage and usage procedures that would improve perceived quality. To achieve improvements, the representatives cleaned and, if necessary, serviced the retailers' equipment; provided free product samples; supplied free accessories; and/or recommended changes in the retailers' current procedures. Before leaving, the representatives demonstrated the improvements by comparing the output produced by the retailer at the beginning of the visit with that produced at the end of the visit. In the month after their visits the representatives followed up with telephone calls and, if appropriate, supplied retailers with additional literature and accessories.

As a record of each visit, the management representatives completed a brief log that summarized the actions they had taken and the retailers' responses. Analysis of these logs indicated that 96% of the retailers visited were supplied with

Table 2
CATEGORIES OF NEEDS

Targeted Needs	*Ancillary Needs*	*Distinct Needs*
Can vary size	Anyone can use	Compact and portable
Color	Can switch types without waste	Equipment looks serious and professional
Finished product has no defects	Equipment and supplies always available	Equipment is durable
Sharp	Manufacturer fixes problems quickly	Manufacturer provides advertising support
Time	Manufacturer that I can deal with easily	Price is reasonable
	Quick to use and trouble free	
	What you see is what you get	

free product, literature, and/or accessories and that in almost every case the retailers acknowledged that there was a dramatic improvement in the perceived quality of the manufactured product they produced. They had not expected a visit from a representative of KemTek but almost universally appreciated the visit.

In the months following the site visits by the management representatives, retailers might or might not have been given a higher than normal level of attention from KemTek's sales representatives. It was not feasible for KemTek to record whether such visits took place, which retailers were visited, or the content of the visits. Because the sales representatives were free to make their own decisions, we assume that they made such visits only if visits did not divert them from other activities they perceived as more productive.

Spain

KemTek intended that the intervention in Spain would yield the same results as in the United States; however, the details would be adapted to the Spanish market. Unlike employees in the United States, the KemTek employees in Spain were not incremental resources but rather a redirection of activities from business as usual to the customer satisfaction improvement program. As a result, the Spanish employees were given more freedom in implementing the intervention. They chose a program that included three site visits to each retailer by local sales representatives of KemTek. The series of visits was positioned as a training program, and retailers were promised a Gold Seal Accreditation upon completion of the visits. Before the program, the representatives received one day of technical training and were accompanied for two days on site visits by expert technicians sent specifically for this purpose from the corporate office. On the first visit the representatives asked for a product sample to assess initial quality. They then presented the goals of the Gold Seal program; gave training on the use and storage of KemTek's product; cleaned and replaced equipment; recommended, sold, and/or installed additional accessories; and finished by asking for another product sample to compare the quality improvement. On the second visit, the representatives installed a checklist that summarized and reiterated their previous advice and followed up on any previous service requests. On the third visit, Gold Seal Accreditation certificates were issued, and any additional service or training requests were satisfied. Approximately 75% of the retailers in the Spanish treatment regions participated in the program. The remaining retailers either could not be located by the implementation team or refused to participate.

ANALYSIS OF THE RELATED QUASI-EXPERIMENTS

The task force designed the interventions in both the United States and Spain to improve satisfaction with respect to the targeted needs. It predicted that improvements on the targeted needs would enhance overall satisfaction and, in turn, lead to more long-term profit for KemTek. Our first analyses test KemTek's predictions. Because this is a quasi-experiment (without random assignment of customers to treatment groups), we must understand the baseline satisfaction in each treatment group. Therefore, we first compare pretest satisfaction of retailers that received the experimental treatment with satisfaction of retailers that did not. The averages of the pretest satisfaction measures in the treatment and control regions are summarized in Table 3. It is evident that in Spain retailers in the treatment cities generally reported higher levels of pretest satisfaction than did retailers in the control cities. It appears that the groups are not equivalent in pretest satisfaction, despite the relatively large sample sizes and KemTek's efforts to identify roughly equivalent groups.

These differences between pretest satisfaction levels in the treatment and control regions suggest that the regions have not been subject to identical histories. We control for these differences in pretest satisfaction of the customer needs with the standard pretest–posttest analysis described subsequently. This analysis ensures that we do not misattribute ex ante differences in the groups to the effects of the treatment.

In addition, the existence of prior differences cautions that differences might persist. Thus, we must consider controls to correct for any potential continued, unobserved ecological changes that might affect the change in satisfaction between the pretest and posttest. This is a serious issue in Spain because KemTek was concerned that unobserved actions by the competitor would lower retailers' perceived satisfaction with all needs. Although KemTek believed that these changes would be the same in the treatment and control cities, the firm did not have the ability to collect detailed information about competitive activity, so unobserved effects were a real possibility. Fortunately, KemTek's extensive data collection gave us the means to control for potential unobserved impacts on customer needs. We do this by using changes in the distinct needs as a nonequivalent dependent variables control (Cook and Campbell 1979, p. 261).

We first develop a model to predict what posttest satisfaction with the customer needs would have been in the absence of an intervention. We develop this model by estimating the following equation using the responses to the five distinct needs, where i indexes individual respondents and n indexes the needs. This equation simultaneously controls for three effects: an individual-specific effect, a need-specific effect, and a pretest measure effect.

$$(1)\ \text{posttest satisfaction}_{in} = \alpha_i + \beta_{1i}\,\text{average pretest satisfaction}_n + \beta_{2i}\,\text{pretest satisfaction}_{in} + \text{error}.$$

The individual-specific effect accounts for heterogeneity in customers' reactions to the scales. (We used the same scale format for all needs and for overall satisfaction.) For each customer, i, we allow a mean bias, α_i, to account for any yea-saying or nay-saying tendency on the part of respondents (Greenleaf 1992). This parameter also controls for halo effects.

The need-specific effect accounts for, on average, some needs being satisfied better than others. We use the average pretest rating of need n. We call this variable average pretest satisfaction$_n$ and allow its contribution to posttest satisfaction to vary by respondent. On the basis of Table 3, we use separate averages for the test and control groups.

The final control variable is the individual respondents' pretest ratings. Because these ratings are measured with error, we must account for their reliability (Caporaso and Roos 1973; Silk 1994). Indeed, if the pretest ratings were the only

Table 3
AVERAGE PRETEST SATISFACTION

	Spain		*United States*	
	Treatment	*Control*	*Treatment*	*Control*
Overall satisfaction	7.91	7.36**	6.80	7.23
Targeted needs				
Can vary in size	3.07	4.66**	5.47	6.62**
Color	7.06	6.88	6.84	6.95
Finished product has no defects	7.63	6.81**	7.10	7.18
Sharp	7.75	7.21*	7.18	7.52
Time	8.96	7.53**	7.56	7.92
Ancillary needs				
Anyone can use	7.96	7.54	7.75	7.73
Can switch types without waste	6.15	7.37**	7.06	6.73
Compact and portable	8.14	7.38**	6.77	7.02
Equipment and supplies always available	8.81	7.74**	8.39	7.91
Manufacturer fixes problems quickly	7.52	7.32	7.63	7.29
Manufacturer that I can deal with easily	7.57	7.79	7.49	7.24
Quick to use and trouble free	8.39	8.02	7.82	7.84
What you see is what you get	8.30	7.52**	7.35	7.36
Distinct needs				
Compact and portable	8.14	7.38*	6.77	7.02
Equipment looks serious and professional	8.11	7.89	6.87	7.22
Equipment is durable	7.27	7.10	7.69	7.44
Manufacturer provides advertising support	7.61	7.26	6.17	6.30
Price is reasonable	6.93	6.51	6.93	6.86
Sample size	139	188	99	125

*The treatment and control averages (in the same country) are significantly different at the .05 level (two-tailed test).
**The treatment and control averages (in the same country) are significantly different at the .01 level (two-tailed test).

variable in the model and we expected no change in the true ratings, the regression coefficient would estimate the reliability.[2] We allow the reliability, β_{2i}, to be heterogeneous.

Although simply comparing the change in satisfaction (posttest satisfaction – pretest satisfaction) is intuitively appealing, the reliability arguments alone argue for a more general model with $\beta_{2i} \neq 1$. In addition, two other arguments suggest that we allow β_{2i} to be estimated. If satisfaction changes over time, we expect current satisfaction to reflect prior satisfaction updated by recent experiences. The coefficient β_{2i} would also reflect the weight assigned to prior satisfaction.[3] Furthermore, pretest satisfaction levels may affect the ability to implement improvements: $\beta_{2i} < 1$ could be interpreted as an indicator that it is harder to improve satisfaction when retailers are already satisfied—in other words, β_{2i} is a correction for scale saturation. All three arguments suggest that $\beta_{2i} < 1$. The estimates from our data turn out to be within this range.

We estimate Equation 1 using data only from the distinct needs, which should be unaffected by the intervention. The data include the individual ratings (327 respondents in Spain and 224 in the United States) on each of the five distinct needs. This is logically equivalent to estimating a separate three-parameter model for each respondent on the basis of observations of the five distinct needs. For the Spanish data this implies a model with 1653 parameters estimated with 2755 observations.[4] We used a Chow (1960) test to compare the fit of this model with a parsimonious model that estimated aggregate coefficients rather than individual coefficients for each respondent (Equation 1 without the i subscripts on α_i, β_{1i}, and β_{2i}). The Chow test rejected the parsimonious model in both Spain and the United States ($p < .01$).

The coefficients estimated for each respondent were used to predict posttest satisfaction for the targeted needs, the ancillary needs, and overall satisfaction. We then test our prediction that satisfaction with the targeted needs will be higher among customers in the treatment cities by comparing observed measures with those predicted by Equation 1. If KemTek's intervention had a measurable effect, the relative effect (observed satisfaction minus satisfaction predicted by pretest measures) should be significantly larger in the treatment regions than in the control regions.

[2]To illustrate this, set up the regression equation $x_2 = \alpha + \beta x_1 +$ error, where both x_1 and x_2 are measured with error. The coefficient, β, is then an estimate of the true variance divided by the total variance.

[3]Bolton and Drew (1991) offer a similar argument and note that this is consistent with a Bayesian framework in which customers use both current and prior information (see also Boulding et al. 1993). Further support for this approach can be found in Bolton and Drew (1991), Caporaso and Roos (1973), and Silk (1994).

[4]Although the individual estimates of posttest satisfaction might have high variance, the estimates of average posttest satisfaction are compared across large sample sizes (test versus control in the United States and Spain) and have many degrees of freedom. This approach is not unlike that used in conjoint analysis when separate response functions are estimated for each respondent, but predictions are based on simulators that aggregate across all respondents. For example, see Green and Srinivasan (1990) and Wittink and Cattin (1989). To test the sensitivity to degrees of freedom, we also estimated a model based on the seven ancillary and five distinct needs. Such models have 12 observations and nine degrees of freedom for each respondent. Significance levels changed slightly, but the results were qualitatively similar.

Residual Satisfaction

Before presenting the results, we introduce one more construct, residual satisfaction. On the basis of KemTek's experience and the voice-of-the-customer theory, KemTek assumed that overall satisfaction could be decomposed on the basis of satisfaction of customer needs (see the review in Griffin and Hauser 1993). KemTek went to considerable effort to ensure an exhaustive list of customer needs. We define residual satisfaction to test KemTek's assumption. We begin by approximating the customer need decomposition with a linear model.

$$(2)\qquad \text{overall posttest satisfaction}_i = \theta_1 + \theta_2\Sigma_n w_{in}\text{posttest satisfaction}_{in} + \text{error}.$$

The terms θ_1 and θ_2 are estimated using ordinary least squares (OLS). The w_{in} is a weight ascribed to each of the 17 needs. For robustness we considered three different methods for determining the weight to ascribe to each need (determining the w_{in} terms): equal weights, stated weights, and revealed weights. The equal weights method attributes the same weight to each need (Einhorn and Hogarth 1975). The stated weights method uses retailers' responses to the importance questions in the pretest and posttest measurement waves. The revealed weights method uses OLS coefficients in which overall satisfaction is regressed on all 17 needs. For parsimony and ease of exposition we focus on the equal weights model. Similar results were obtained using the stated weights model. Weights in the revealed weights model cannot be estimated reliably because of collinearity between the needs. (See also comparisons in Griffin and Hauser 1993.)

Residual satisfaction then is that portion of overall satisfaction that cannot be explained with the measures of satisfaction for the customer needs. That is,

$$(3)\qquad \text{residual satisfaction}_i = \text{overall posttest satisfaction}_i - \theta_1 - \theta_2\Sigma_n w_{in}\text{posttest satisfaction}_{in}.$$

We check whether residual satisfaction is larger (or smaller) in the treatment regions than in the control regions. To control for unobserved ecological changes, we again use the nonequivalent dependent variables control. In particular, we separately calculate both an observed and a predicted measure of residual satisfaction. The observed measure of residual satisfaction represents the residuals from estimating Equation 2 using the observed measures of overall satisfaction and satisfaction with the individual needs. We calculate the predicted measure of residual satisfaction using a two-step process. We first use Equation 1 to derive estimates of overall satisfaction and satisfaction with the customers' needs. We then reestimate Equation 2 with these predictions and retain the residuals from this estimate as our measure of predicted residual satisfaction.[5] If variance in the 17 customer needs is sufficient to explain systematic variance in overall satisfaction, the difference between observed and predicted residual satisfaction should not be significant in the test region and in the control region.

[5]To control for the possibility that the intervention changed θ_1 and θ_2, we calibrate Equation 2 separately using the actual and predicted measures.

Table 4

DIFFERENCES IN POSTTEST SATISFACTION BETWEEN CONTROL REGIONS AND TEST REGIONS CONTROLLING FOR INDIVIDUAL DIFFERENCES (EQUATION 1)

	Spain	*United States*
Overall satisfaction	.57*	.15
Targeted needs	.60*	.38*
Ancillary needs	.48**	.00
Residual satisfaction	.25**	.17
Number of respondents		
Treatment group	133	96
Control group	182	124

*Adjusted posttest satisfaction is significantly larger in the treatment group than in the control group (in the same country) at the .05 level (one-tailed test).

**Adjusted posttest satisfaction is significantly larger in the treatment group than in the control group (in the same country) at the .01 level (one-tailed test).

Notes: The data in the table correspond to the differences in average adjusted posttest satisfaction between the treatment and control groups. Adjusted posttest satisfaction$_{in}$ = posttest satisfaction$_{in}$ – α_i – β_{i1}average pretest satisfaction$_n$ – β_{i2}pretest satisfaction$_{in}$ calibrated using satisfaction with the five distinct needs.

The sample size for residual satisfaction and overall satisfaction is the number of respondents. Sample sizes for the targeted and ancillary needs comparisons are five and seven times larger, respectively.

RESULTS

Table 4 summarizes the net impact of the U.S. and Spanish interventions.[6] The distinct needs act as a control and therefore are not included in Table 4. The results support our prediction that satisfaction with the targeted needs will be higher in the treatment cities. The quality improvements yielded enduring and measurable improvements in customer satisfaction with the targeted needs. These results are comforting. A carefully designed and implemented customer satisfaction improvement intervention could yield positive results. Despite the lack of significance in the United States for overall satisfaction, KemTek considered all the data and, combined with managerial judgment, believed that both the U.S. and Spanish interventions achieved their objectives.

KemTek continued its customer satisfaction initiatives. It is beyond the scope of this article (and proprietary to KemTek) to discuss the details necessary to estimate whether the increased revenues justified the interventions' costs. We can only say that, today, customer satisfaction is an important criterion by which executives at KemTek are evaluated.

However, there are two surprises (see Table 4). First, the results also suggest that there were differences between the U.S. and Spanish interventions, even though KemTek believed ex ante that they would be equivalent. It appears that the scope of the intervention in Spain was much broader than that in the United States. Although the intervention in

[6]The sample size is slightly smaller for Table 4 than for Table 3 for technical reasons. That is, for some individual respondents there is not sufficient variance in the distinct needs to identify the three individual-specific parameters. For example, Equation 1 becomes overspecified if a respondent gave the same pretest satisfaction response for all 5 distinct needs. Fewer observations are lost in a model that uses the 12 needs (distinct and ancillary). Such a model gives qualitatively similar results.

the United States appeared to affect satisfaction with the targeted needs, it did not appear to affect the ancillary needs. The change in overall satisfaction had a positive sign in the United States, but the change was not significant. In contrast, the Spanish intervention appeared to increase satisfaction with the ancillary needs and overall satisfaction. Second, in Spain, the changes in the customer needs did not appear to be sufficient to explain the change in overall satisfaction. There was also a significant impact on residual satisfaction.

The data in Table 4 can only highlight the surprises, not explain them. However, because we were involved from the beginning and have access to the paper trail, we use our experience to conjecture on (1) the cause of the apparent difference in scope of the two interventions and (2) the significant effect in Spain on residual satisfaction. We address each of these in turn. We then discuss another important lesson from the quasi-experiments—the practical importance of the nonequivalent dependent variables design, which was necessary to identify the significance of the interventions.

Surprise 1: The Variation in the Scope of the Intervention Between the United States and Spain

The U.S. intervention was implemented by management volunteers who had considerable technical expertise but limited previous interaction with retailers in this market. The absence of market knowledge made these management volunteers dependent on the task force and the local U.S. management for guidance in conducting their field visits. This guidance was provided in a formal training session and through detailed scripts and procedure manuals that the task force reviewed and helped design. The influence of the task force and the volunteers' need for guidance ensured that the intervention was closely focused on the targeted needs.

In contrast, the Spanish interventions were implemented by KemTek's local sales representatives, who had extensive market experience but limited technical expertise. These representatives received some technical training but little other guidance, and their activities were subjected to less review and control by the task force. In the absence of that control, the Spanish representatives may have diverted their efforts from the specific activities proposed by the task force to other improvements suggested by their knowledge of the market. Lessening the task force's control appears to have yielded more wide-ranging improvements, without compromising the impact on the targeted needs. To the extent that this holds up in other quasi-experiments, this result argues for a policy of allowing resources to be allocated by the parties who are best informed about their use.

Other differences between the Spanish and U.S. interventions can be interpreted as an outcome of empowering the Spanish representatives. For example, the Spanish representatives chose to make three separate visits to the treatment retailers, whereas their U.S. counterparts were instructed to make just a single visit. This difference provided the Spanish representatives with more time to satisfy a broader range of retailer needs. If the same level of control had been exercised over the Spanish representatives as was exercised over their U.S. counterparts, the Spanish intervention probably would also have been a single visit.

Although this ex post analysis highlights the differences between the U.S. and Spanish interventions, KemTek's management did not have the luxury of this data when it designed the interventions. KemTek is an experienced multinational firm operating in many geographic markets. It decided to control carefully the U.S. management volunteers and to impose much less control in Spain. Ex ante, KemTek did not believe the empowerment of the Spanish representatives was a major difference. Ex post, we now realize it may have been a major difference. Although the quasi-experiments by themselves cannot rule out other differences between the United States and Spain, such as culture, language, and the presence of competition, empowerment survives as an attractive explanation. For example, experienced KemTek managers did not believe culture and language caused the difference. The presence of competition differed between the quasi-experiments, but KemTek's hypothesis was that competition would decrease the impact of the intervention, not make it more wide-ranging.

We believe that this potential evidence for empowerment is a major practical lesson, both for the manner in which global marketers approach their markets and as evidence of the efficiency of trusting in local marketing knowledge. At a minimum it is an interesting hypothesis worth further testing.

Surprise 2: A Significant Increase in Residual Satisfaction in Spain

The increase in residual satisfaction in Spain offers evidence that the improvement in overall satisfaction due to the intervention cannot be fully explained by the changes in satisfaction with the 17 measured needs. More generally, this suggests that overall satisfaction in Spain is affected by factors other than the 17 measured needs. This result occurred despite the considerable resources that KemTek invested to ensure that no retailer needs were omitted. Professionals and managers with extensive experience in the relevant products and markets used state-of-the-art methods.

We can estimate the likelihood of missing customer needs by using Griffin and Hauser's (1993) beta-binomial model. Their model suggests that 99% of the product and service delivery needs were uncovered by the 38 merged interviews (the relevant model is the improved questioning method; Griffin and Hauser 1993, p. 10). Even if we limit the analysis to the 20 Spanish interviews, the model suggests that 98% of the product and service needs were uncovered. It is unlikely that KemTek missed a sufficient number of retailer needs to explain the significant increase in residual satisfaction. It is more likely that the intervention in Spain affected constructs that do not fit Griffin and Hauser's (1993, p. 4) definition of "a description, in the customer's own words, of the benefit fulfilled by the product or service experience."

Our residual satisfaction estimation procedure controlled for changes in the importance of the customer needs; thus, we also reject that potential explanation for the observed effect. In general, residual satisfaction might be due to nonlinearities in the relationship between needs and overall satisfaction (Mittal, Ross, and Baldasare 1998). We do not think that is the case here because (1) we tested for nonlinearities and did not find them and (2) linear models have fitted well in the past.[7]

[7]We considered log transformations and the introduction of quadratic terms. For a discussion of the robustness of linear models, see Griffin and Hauser (1993).

However, the hypothesis that residual satisfaction in Spain reflects one or more unmeasured determinants of overall satisfaction is consistent with our previous arguments that the Spanish intervention was broader in scope than the U.S. intervention. The Spanish representatives may have found a way to enhance overall satisfaction directly rather than through the 17 needs—a way not anticipated by the task force. We offer three hypotheses to suggest further research.

Labeling and self-perception. The Spanish representatives chose to give retailers in Spain that completed the three-step training and service program a Gold Seal Accreditation, which certified the retailers' participation. Retailers in both the United States and Spain were given positive reinforcement when they were shown how much better they could use the equipment after receiving training. However, the accreditation was used only in Spain. The labeling literature suggests that the very act of certifying successful completion may have influenced retailers' perceptions. Labeling a retailer as the type of company that would tie itself to KemTek (certification) might lead to behavior and beliefs consistent with the label (Allen 1982; Allen and Dillon 1983). Because retailers were KemTek customers, the label is consistent with their self-schema and therefore more likely to be salient (Tybout and Yalch 1980). In addition, the Spanish retailers' investment of their own time in the training may have led to a perception that it was worthwhile to link themselves to KemTek (Bem 1972; Folkes and Kiesler 1991).

Commitment and trust. The Spanish representatives chose to make three visits rather than just one. (There was a follow-up in the United States, but it was only by telephone.) On each subsequent visit, the representatives responded to requests made on the previous visit. This might signal the desire to invest in a durable relationship of shared interests, which could lead to commitment and trust (Dwyer, Schurr, and Oh 1987). This hypothesis is consistent with the work of Morgan and Hunt (1994), who propose commitment and trust as important determinants of successful channel relationships, and is in apparent accordance with predictions from the economic literature on repeated games (Abreu 1988; Axelrod 1984).[8]

Spain Versus the United States. Finally, the effects occurred in Spain but not in the United States. Although KemTek did not believe that language or culture was the determinant, we cannot rule out the hypothesis that mechanisms of customer satisfaction vary on the basis of language and culture.

It is beyond the scope of this article (and KemTek's data) to test these hypotheses. However, we suggest that such data be collected in future customer satisfaction interventions. Measurement scales exist in the literature for these constructs (e.g., Morgan and Hunt 1994; Sullivan et al. 1981).

[8]Commitment is an "enduring desire to maintain a valued relationship" (Moorman, Zaltman, and Deshpandé 1992, p. 316) and "an implicit or explicit pledge of relational continuity" (Dwyer, Schurr, and Oh 1987, p. 19). Trust is "a willingness to rely on an exchange partner in whom one has confidence" (Moorman, Zaltman, and Deshpandé 1992, p. 315) and "the belief that a party's word or promise is reliable and a party will fulfill his/her obligations in an exchange relationship" (Schurr and Ozanne 1985, p. 940).

Figure 2
UNTREATED CONTROL GROUP DESIGN WITH PRETEST AND POSTTEST

O_{1A}	×	O_{2A}
O_{1B}		O_{2B}

PRACTICAL NEED FOR THE NONEQUIVALENT DEPENDENT VARIABLES CONTROL

KemTek invested in extensive measures to determine whether its customer satisfaction intervention had the desired effect. These measures included both pretest and posttest measures, control groups, and nonequivalent dependent measures. This level of measurement is well beyond that which is typical in industry. From our experience, the most common designs are one-group pretest–posttest designs or, sometimes, posttest-only designs. The limitations of these designs are widely recognized and well understood in the academic literature (e.g., Cook and Campbell 1979, p. 247). Nonetheless, the wide industrial use of such designs might lead to false rejection of customer satisfaction initiatives. For example, in Spain, where satisfaction with all customer needs was generally trending downward (likely due to competitive actions), had we analyzed KemTek's test groups only we would have found either no effect or a negative effect.

The more interesting aspect of KemTek's design was the availability of the nonequivalent dependent variables. These variables, which were clearly not targeted by the intervention, enabled us to control for the otherwise unobservable ecological impacts on all customer needs. To illustrate their impact, we reanalyze the data with the more typical untreated control group design with pretest and posttest illustrated in Figure 2.

With this design we can no longer estimate Equation 1 because the nonequivalent dependent variables are not available. However, we can control for individual scale effects, using the average of overall satisfaction and the responses to the targeted and ancillary need questions (for each respondent). We can also estimate a question effect, using the average response (across all respondents) to each question. The relevant equation then becomes

$$\text{(4)}\quad \text{posttest satisfaction}_{in} = \alpha + \beta_1\text{average pretest satisfaction}_n + \beta_2\text{average pretest satisfaction}_i + \beta_3\text{pretest satisfaction}_{in} + \beta_4\text{intervention}.$$

The results based on this design are summarized in Table 5. In the United States the results are qualitatively similar, perhaps because there was no competitor to cause unobserved ecological changes in all customer needs. However, in Spain, where there was likely significant but unobserved competitive activity, the results change dramatically. There is still a significant impact on overall satisfaction and residual satisfaction, but there was no significant effect on the targeted needs or on the ancillary needs (in fact, the signs are negative). Without the nonequivalent dependent variable controls, the analysis in Table 5 might have falsely rejected

Table 5
ESTIMATED IMPACT OF THE TREATMENT ON POSTTEST SATISFACTION WITHOUT CONTROLLING FOR CHANGES IN DISTINCT NEEDS

Variables	*Spain*	*United States*
Overall satisfaction	.46*	.28
Targeted needs	–.03	.48**
Ancillary needs	–.15	.15
Residual satisfaction	.60**	.14
Number of respondents		
Treatment groups	139	99
Control groups	188	125

*The β_4 coefficient is significantly larger than zero at the .05 level (one-tailed test).

**The β_4 coefficient is significantly larger than zero at the .01 level (one-tailed test).

Notes: The data in the table describe the β_4 coefficient from the following model: posttest satisfaction$_{in}$ = α + $\beta_1$average pretest satisfaction$_n$ + $\beta_2$average pretest satisfaction$_i$ + $\beta_3$pretest satisfaction$_{in}$ + $\beta_4$intervention estimated on the treatment and control groups in each country. For overall satisfaction and residual satisfaction, α and β_1 cannot be estimated independently; thus, β_1 is restricted to equal zero.

The sample sizes for the targeted and ancillary needs models are five and seven times larger than the number of respondents (respectively).

the ability of the customer satisfaction intervention to affect the targeted needs. It is also possible that industry would consider an even simpler model that does not account for the reliability of the measures. One such model might simply examine the differences in the means between the pretest and posttest measures. When we examined such a model, it also estimated a significant increase in the targeted needs in the United States and a nonsignificant decrease in the targeted needs in Spain.

POTENTIAL LIMITATIONS

Although KemTek collected more data than is normal in industrial settings, we caution the reader that the two interventions were quasi-experiments rather than fully controlled experiments. KemTek wanted to understand the results of the interventions, but it had to balance this goal with its fiduciary responsibility of earning profit in these markets. As a result, the U.S. and Spanish interventions differ on more than one dimension. We have done our best to interpret these differences in the light of our knowledge of the interventions and KemTek's knowledge of the markets, but the natural limits of quasi-experiments remain.

In addition, although our experience and that of KemTek suggest that the distinct needs were appropriate as nonequivalent dependent variable controls, it is always possible that there was some small impact on the distinct needs that was due to the intervention. For example, retailer perceptions of satisfaction with price may have changed more in the treatment regions than in the control regions, even though the actual measures—say, the price of the product—remained unchanged. An alternative explanation that attempts to explain such a change might be that perceptions of the distinct needs required maintenance effort, and the Spanish representatives diverted efforts from the distinct needs toward the targeted needs. Although neither KemTek nor we believe this was the explanation, we cannot rule it out completely. Our experiences suggest that it is more likely that competitive entry targeted all needs (targeted, ancillary, and distinct) and that, without the intervention, all needs would have been lowered. Even if we accept this alternative explanation that the effect of the intervention was only relative, KemTek still considered the intervention successful. The targeted needs were chosen because they were most important to customers, but there is evidence that overall satisfaction increased as well. It increases when the distinct needs act as controls (Table 4) and, as indicated in Table 5, it increases even when the distinct needs are not used as controls.

CONCLUSIONS

We reported on the results of controlled, longitudinal field studies in two countries, in which a sophisticated, high-technology firm used state-of-the-art marketing research and quality tools to design and implement a customer satisfaction improvement program. The results confirm the basic premise that it is possible to implement quality improvements that yield enduring and measurable improvements in customers' perceptions of satisfaction. This finding is particularly notable because of the delay between the intervention and posttest measures and the rather targeted nature of the treatment. The experimental treatment focused on providing training to customers in the use and storage of a business-to-business product. No changes were made to the price or the production, distribution, or sales systems. The intervention was successful in a country in which the firm enjoyed an effective monopoly and in a representative country in which the firm faced a strong competitive entrant.

Besides demonstrating that a carefully designed customer satisfaction intervention could be successful in a field setting, the matched quasi-experiments highlight three interesting lessons. First, the more broad-ranging impact in Spain suggests that firms can combine careful central planning (voice of the customer, house of quality, interfunctional task force) with a strategy that empowers employees to adapt interventions to local market conditions. Second, the ability of the Spanish representatives to effect a significant improvement in residual satisfaction suggests the need to understand further those aspects of overall satisfaction that can be affected independently of the satisfaction of the customer needs. Third, the Spanish analyses, compared with the U.S. analyses, suggest that the popular press and industry might falsely reject customer satisfaction initiatives because they rely on insufficient controls to evaluate the initiatives properly. KemTek collected data that were well beyond industry norms. Perhaps those norms need to be rethought.

REFERENCES

Abreu, Dilip (1988), "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56 (March), 383–96.

Allen, Chris T. (1982), "Self-Perception Based Strategies for Stimulating Energy Conservation," *Journal of Consumer Research*, 8 (March), 381–90.

——— and William R. Dillon (1983), "Self-Perception Development and Consumer Choice Criteria: Is There a Linkage?" in *Advances in Consumer Research*, Vol. 10, Richard P. Bagozzi and Alice M. Tybout, eds. Ann Arbor, MI: Association for Consumer Research, 45–50.

Anderson, Eugene W., Claes Fornell, and Donald R. Lehmann (1994), "Customer Satisfaction, Market Share, and Profitability: Findings from Sweden," *Journal of Marketing*, 58 (July), 53–66.

——— and Mary W. Sullivan (1993), "The Antecedents and Consequences of Customer Satisfaction for Firms," *Marketing Science*, 12 (Spring), 125–43.

Axelrod, Robert (1984), *The Evolution of Cooperation*. New York: Basic Books.

Beckwith, Neil E. and Donald R. Lehmann (1976), "Halo Effects in Multiattribute Attitude Models: An Appraisal of Some Unresolved Issues," *Journal of Marketing Research*, 13 (November), 418–21.

——— and ——— (1975), "The Importance of Halo Effects in Multi-Attribute Models," *Journal of Marketing Research*, 12 (August), 265–75.

Bem, Daryl (1972), "Self Perception Theory," in *Advances in Experimental Social Psychology*, Vol. 6, Leonard Berkowitz, ed. New York: Academic Press, 199–218.

Bolton, Ruth N. and James H. Drew (1991), "A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes," *Journal of Marketing*, 55 (January), 1–9.

Boulding, William, Ajay Kalra, Richard Staelin, and Valarie A. Zeithaml (1993), "A Dynamic Process Model of Service Quality: From Expectations to Behavioral Intentions," *Journal of Marketing Research*, 30 (February), 7–27.

Caporaso, James A. and Leslie L. Roos Jr. (1973), *Quasi-Experimental Approaches: Testing Theory and Evaluating Policy*. Evanston, IL: Northwestern University Press.

Chow, Gregory C. (1960), "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," *Econometrica*, 28 (July), 591–605.

Cook, Thomas D. and Donald T. Campbell (1979), *Quasi-Experimentation: Design and Analysis Issues For Field Settings*. Boston: Houghton Mifflin.

Cronin, J. Joseph and Steven A. Taylor (1994), "SERVPREF Versus SERVQUAL: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality," *Journal of Marketing*, 58 (January), 125–31.

Crosby, Lawrence A. and Nancy Stephens (1987), "Effects of Relationship Marketing on Satisfaction, Retention, and Prices in the Life Insurance Industry," *Journal of Marketing Research*, 24 (November), 404–11.

Dwyer, F. Robert, Paul H. Schurr, and Sejo Oh (1987), "Developing the Buyer–Seller Relationships," *Journal of Marketing*, 51 (April), 11–27.

Einhorn, Hillel J. and Robin M. Hogarth (1975), "Unit Weighting Schemes for Decision Making," *Organizational Behavior and Human Performance*, 13, 171–92.

Folkes, Valerie S. and Tina Kiesler (1991), "Social Cognition: Consumers' Inferences About the Self and Others," in *Handbook of Consumer Behavior*, Thomas S. Robertson and Harold H. Kassarjian, eds. Englewood Cliffs, NJ: Prentice-Hall, 281–315.

Fornell, Claes (1992), "A National Customer Satisfaction Barometer," *Journal of Marketing*, 56 (January), 6–21.

Fournier, Susan and David G. Mick (1999), "Rediscovering Satisfaction," *Marketing Science Institute Working Paper No. 99-092*. Cambridge, MA: Marketing Science Institute.

Green, Paul E. and V. Srinivasan (1990), "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Marketing*, 54 (October), 3–19.

Greenleaf, Eric A. (1992), "Improving Rating Scale Measures by Detecting and Correcting Bias Components in Some Response Scales," *Journal of Marketing Research*, 29 (May), 176–88.

Griffin, Abbie and John R. Hauser (1993), "The Voice of the Customer," *Marketing Science*, 12 (Winter), 1–25.

Hauser, John R., Duncan I. Simester, and Birger Wernerfelt (1994), "Customer Satisfaction Incentives," *Marketing Science*, 13 (Fall), 327–50.

———, ———, and ——— (1996), "Internal Customers and Internal Suppliers," *Journal of Marketing Research*, 33 (3), 268–80.

———, ———, and ——— (1997), "Side Payments in Marketing," *Marketing Science*, 16 (3), 246–55.

Howe, Roger J., Dee Gaeddert, and Maynard A. Howe (1995), *Quality on Trial: Bringing Bottom-Line Accountability to the Quality Effort*. New York: McGraw-Hill.

Mittal, Vikas, William T. Ross Jr., and Patrick M. Baldasare (1998), "The Asymmetric Impact of Negative and Positive Attribute-Level Performance on Overall Satisfaction and Repurchase Intentions," *Journal of Marketing*, 62 (January), 33–47.

Moorman, Christine, Gerald Zaltman, and Rohit Deshpandé (1992), "Relationships Between Providers and Users of Marketing Research: The Dynamics of Trust Within and Between Organizations," *Journal of Marketing Research*, 29 (August), 314–29.

Morgan, Robert M. and Shelby D. Hunt (1994), "The Commitment-Trust Theory of Relationship Marketing," *Journal of Marketing*, 58 (July), 20–38.

Rust, Roland T., Anthony J. Zahorik, and Timothy L. Keiningham (1995), "Return on Quality (ROQ): Making Service Quality Financially Accountable," *Journal of Marketing*, 59 (April), 58–70.

Schurr, Paul H. and Julie L. Ozanne (1985), "Influences on Exchange Processes: Buyers' Preconceptions of a Seller's Trustworthiness and Bargaining Toughness," *Journal of Consumer Research*, 11 (March), 939–53.

Silk, Alvin J. (1994), "Notes on Reliability and Attitude Measurement," *Harvard Business School Case Notes,* (April), 9-594-087.

Sullivan, Jeremiah, Richard B. Peterson, Naoki Kameda, and Justin Shimada (1981), "The Relationship Between Conflict Resolution Approaches and Trust—A Cross Cultural Study," *Academy of Management Journal*, 24 (4), 803–15.

Tybout, Alice M. and Richard R. Yalch (1980), "The Effect of Experience: A Matter of Salience," *Journal of Consumer Research*, 6 (March), 406–13.

Wittink, Dick R. and Phillippe Cattin (1989), "Commercial Use of Conjoint Analysis: An Update," *Journal of Marketing*, 53 (July), 91–96.

Yi, Youjae (1990), "A Critical Review of Consumer Satisfaction," in *Review in Marketing 1990*, Valerie A. Zeithaml, ed. Chicago: American Marketing Association, 68–123.