USING CUSTOMER SATISFACTION DATA TO IMPROVE SERVICE OPERATIONS

James H. Drew, GTE Laboratories, Inc.

ABSTRACT

In this paper, we relate developments in the measurement of customer satisfaction with advances in the philosophy and techniques of quality improvement. This work lies at the intersection of three distinct parts of a service provider's organizational structure--operations, performance management and quality improvement--and provides a conceptual focus for all three. Consequently, customer satisfaction measurement lies at both the beginning and the end of a customer feedback loop in which satisfaction is related back to operations indicators for service monitoring and service improvements, and these improvements are related forward to customer reaction in terms of stated satisfaction and behavior. We illustrate the steps of this loop with a number of statistical analyses for residential telephone service. Customer satisfaction is related therein to specific operational perceptions and key company-generated internal measures. In our examples, the differential effect of telephone repair times on satisfaction is shown, and related to the success of a change in repair policy and in the likelihood of repair guarantee invocation.

INTRODUCTION

Customer satisfaction (CS) measurement systems have become an extremely important part of many American companies' service strategies. There are several stages of sophistication in how these measurement systems are motivated and implemented, however. Initial CS systems often tend to focus on tracking customer survey ratings over time or benchmarked against competitors' ratings. Customer satisfaction ratings become, in effect, a goal in themselves. Further, this information is often examined in isolation from other indicators of the company's operations and strategies.

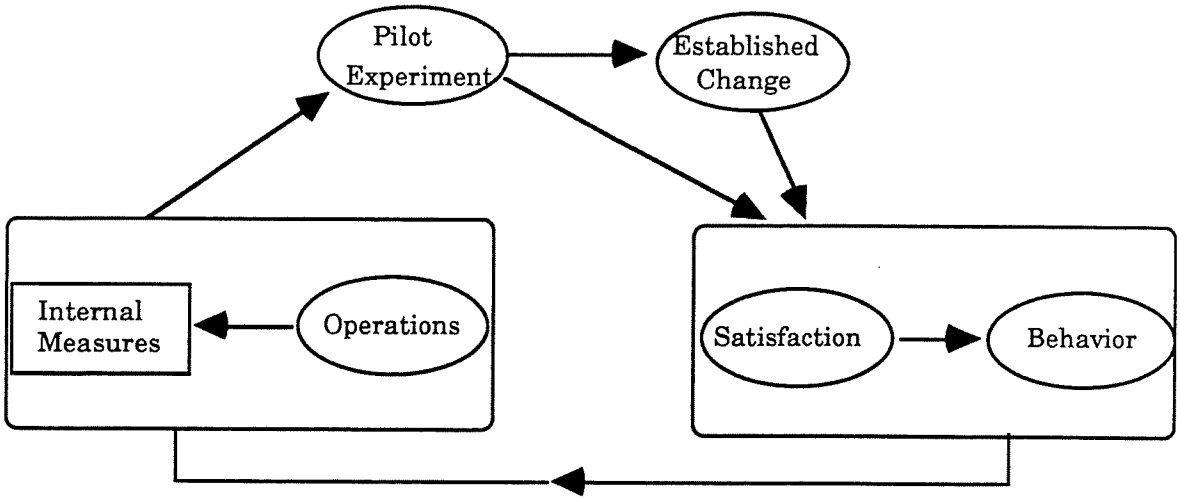
A more sophisticated system would link specific customer evaluations and perceptions to chronic operational characteristics. This linkage may relate internal company measures with customer perceptions, so that, in effect, an internal system is created to detect customer problems. Alternatively, the linkage may be directly between specific company operations and customer evaluations. A qualitative framework for this linkage is the "House of Quality" paradigm of Hauser and Clausing (1988). Specific research that relates service attributes to customer perceptions has been performed for outpatient health services (Neslin, 1983) and engineering (Narasimhan & Sen, 1990), among others.

A third level of sophistication extends the company's interest in customer measurement beyond the operations departments. In accordance with modern Total Quality Management philosophies (see, e.g., Deming 1986), company planning, strategy and operations all begin with the voice of the customer. Among many other things, this central tenet implies that customer satisfaction and evaluation become both the beginning and the end of a feedback loop between the customer, routine operations and service improvements.

Customer attitudes are indicated by ratings on a periodic and extensive satisfaction survey, which are then linked, often in conjunction with other customer and operations intelligence (such as company benchmarking or customer focus groups) with the internal operations of the company to identify any areas for improvement. Internal measures are often an indicator of company operations. Experimental service improvement projects are conceived and piloted in a carefully designed study, whose centerpiece is the measurement of comparative customer reaction and consequent behavior.

In addition, customer satisfaction is a useful construct in the linkage of customer assessments of service to future behavior and therefore to company revenues. Satisfaction is thought to increase customer retention rates (Fornell, 1992), which is well known to be more cost efficient than attracting new customers. Further, several studies find that customer satisfaction is positively related to re-purchase intentions (e.g. Anderson & Sullivan, 1992; Bearden and Teel, 1983). Bolton and Drew's (1992) study of small business customers indicated that perceived service value is positively related to behavioral intentions.

Figure 1



LaBarbera and Mazursky's (1983) longitudinal study supports the role of satisfaction in influencing purchase intentions and behavior.

Note that for a service provider, the centrality of customer satisfaction measurement can be motivated as a logical next step in the evolution of total quality management. One of the hallmarks of the evolution of quality management methods has been to increase its scope from inspection of the finished product, to control of production processes, to the integrated examination of all business functions associated with an offering's provision and maintenance. Consequently, one of the goals of modern quality management is to measure as much of the production process as possible, as locally and concurrently as possible.

This has an important impact on the quality management of services, for by definition, a service is produced and consumed simultaneously. That is, a customer's use of a service is inseparable from the process by which it is generated. Therefore, the measurement of total process quality for a service must include the measurement of its quality in consumption, and it follows that this measurement must originate with the customer. In this sense, note that the upper part of the above diagram, in which the CS-Operations axis is followed by pilot changes which (upon successful trials) become programmatic improvements, is essentially the Plan-Do-Check-Act (PDCA) cycle celebrated in the Quality Management literature (see, e.g. Walton, 1986).

Of course, there are several important aspects of the PDCA cycle and other aspects of Quality Management which are not illustrated in this diagram nor covered in this present work. Many other information sources enter into the planning phase of the cycle. In particular, benchmarking is an important performance and deficiency indicator, both in the sense of surveying the customers of one's competitors and in comparing one's internal processes to those of industry leaders. Further, the activity of setting and meeting process specifications, which is an important part of classical quality assurance, (as well as being central to the service delivery model of Parasuraman *et al*, 1986) is only implicit in the "Operations" box on the left side of the diagram.

To be useful, then the satisfaction construct should be related to the more objective measures of operational attributes and improvements, as well as customer behavior. The qualitative distinction between such subjective and objective sorts of data leads to special problems of data collection and analysis, which we illustrate in subsequent sections for the case of residential telephone service repair.

TRANSLATING CUSTOMER EVALUATIONS INTO OPERATIONAL ATTRIBUTES

It is well known that customers have a

different perspective on the service they receive than does its provider, and the language they use in its description or evaluation may not correspond well with that of operations engineers. It is important, then, to relate overall service evaluations with specific service perceptions. This is the second of the two links between service features and customer perceptions which Holbrook (1981) distinguishes.

Customer evaluations of products and services is often accomplished by sample surveys whose questionnaires should include a large number of items, and particularly, multiple measures of unobservable factors. Their modeling, and potential linkage to other measures (e.g. internal company measures, or operational characteristics) is rather more complicated than the analysis of physical measurements.

Measuring customer affect using different survey items leads directly to the notion that each item measures a set of underlying latent variables with measurement error, whose magnitudes can be estimated through some form of factor analysis. Then, relating customer concepts must be accomplished by structural equation models, in which the relations exist among the latent factors, and not necessarily directly among the observed survey items.

Repair service is an important component of local telephone provision, just as it is for many other products and services. To understand the relationship between repair characteristics and the customer's perception of that service, a random sample of 300 customers with recent repair experiences was surveyed using a questionnaire with items encompassing objective repair characteristics:

DONEPROM - was the repair done when promised?

MORE1 - was the repair requested more than once?

LENGTH - how long did the repair take?

COMPLETE - was completion of the repair verbally confirmed?

as well as more subjective evaluations:

QREPSERV - what was the quality of the entire repair service process?

among others.

An often-used technique for quantifying the effect of repair characteristics on the customer's subjective evaluation is to perform an ordinary least squares regression of QREPSERV on the perceived characteristics MORE1, DONEPROM, LENGTH and COMPLETE. To display the relative sizes of the regression coefficients, the equation below arbitrarily normalizes the coefficient of MORE1 to 1.0:

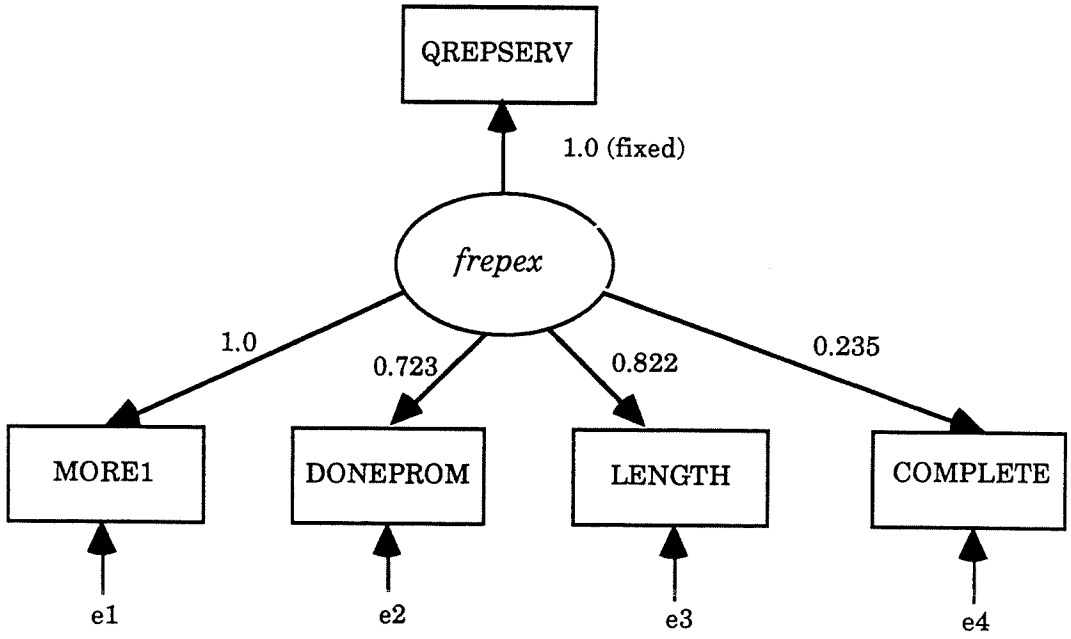
$$\text{QREPSERV} = 1.0\ \text{MORE1} + 0.315\ \text{DONEPROM} + 0.455\ \text{LENGTH} + 0.407\ \text{COMPLETE} + e$$

Two problems with this type of analysis when the explanatory variables are themselves random variables (by virtue of their being measured by customer perception) is that the coefficients are attenuated (see, e.g. Fuller, 1987), and that there is no allowance for the non-zero correlations among these variables.

The figure below shows a simple set of structural equations which address both these problems. A latent factor *frepex* is postulated which generates both QREPSERV and all four explanatory variables, so their correlations are explained by their common origin. As a consequence of our setting the coefficient of QREPSERV to 1.0, *frepex* can be interpreted as a summary of a customer's evaluation of the quality of the repair experience, cleansed of measurement error. Additionally, all five variables are explicitly allowed some measurement errors which, as the results below show, differentially adjust the coefficients from the standard regression model. Finally, the underlying factor is itself allowed to be a random variable, in recognition of its different values for different customers. The figure below shows the model, with named rectangles denoting survey questions, the named oval indicating the unobserved factor *frepex*, and the straight arrows denoting coefficients:

There are substantial practical differences using the more appropriate model. While notification of a repair completion (COMPLETE) has a relatively high value in the regression, the

Figure 2



structural equations show its lower effect in comparison to repair length and the number of repair requests made, as one might intuitively suspect. The relative size of the coefficients of this model suggests the company's need to make the repair as quickly as possible, both to decrease LENGTH and to decrease the likelihood that the customer will call more than once. Note, too, that these attributes are each more important to the customer than whether the repair was completed when promised. This will be seen explicitly below.

RELATING SATISFACTION TO INTERNAL MEASURES: TELEPHONE REPAIR

One of the most serious goals of using data from customer surveys is linking subjective evaluations to company operations. There is often a great need to quantify the relationship between customer perceptions and internal processes; this is the first link between service features and customer assessments described by Holbrook (1981). Managers may wish to prioritize proposed internal improvements in the light of customer needs, while others may wish to anticipate the likely effect of a process change on customer opinion. For these kind of analyses, quantifiable process measures are chosen for the study of their relation to customer ratings.

Although they are not necessarily precise indicators of all aspects of an internal process, managers often find quantified internal measures to be substantially more manageable than direct customer data. Not all internal measures can be usefully related to customer opinion data. Often internal data are generated or collected at a level which does not correspond to the data produced by a given customer. Measures far removed from the actual provision of service to the customer, such as employee absenteeism rates, usually have too subtle an effect on customers for any quantitative analysis to show a linkage. Even when internal measures are gathered for process aspects close to service consumption, they may be aggregated, and not readily available at the level of the individual customer. It is not difficult to imagine plausible situations where correlations based on aggregate data can be very different, and even of opposite sign, from the analogous analysis performed on individual-level data. Further, variations among individuals tend to attenuate correlations or regression coefficients, so aggregate data can only be used with considerable caution, and when there is no feasible individual-level alternative.

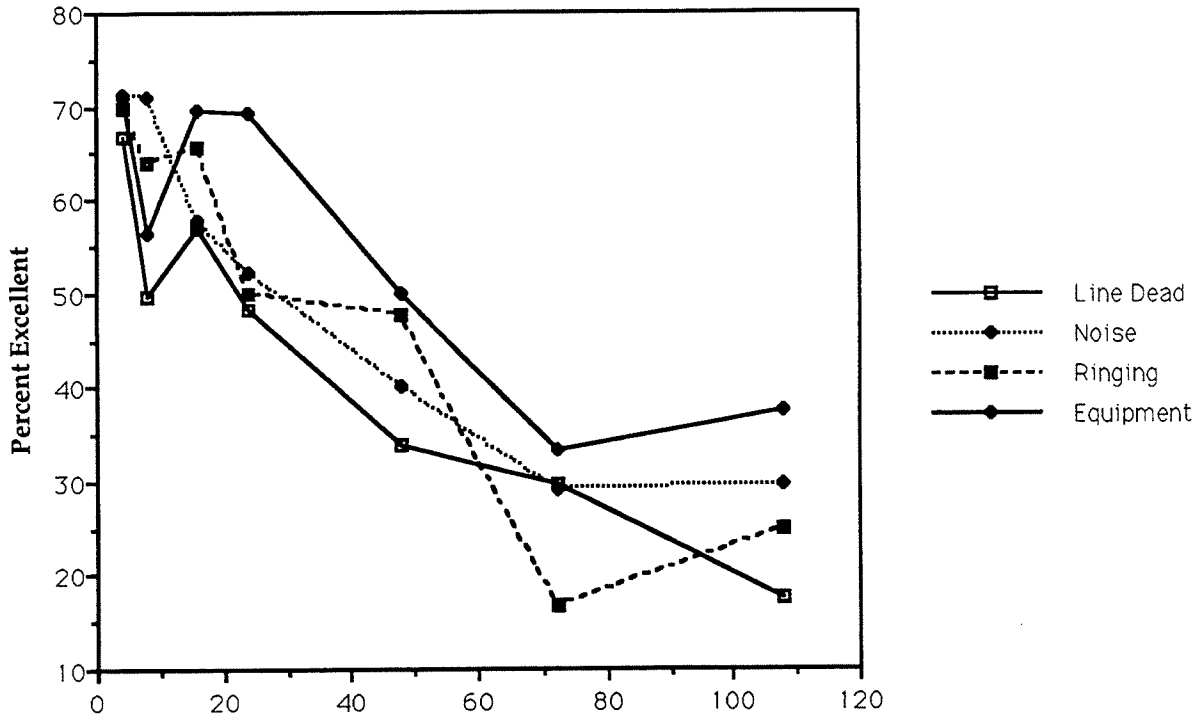
The example of the previous section was based

on a study of 300 residential telephone customers experiencing a repair. Internal measures were recorded for each customer's repair, including its symptoms, duration and cause. As mentioned above, each customer was surveyed shortly after the completion of each repair to obtain his or her perceptions and evaluations of the experience. Thus, individual-level data are available to link the internal and customer-based measures.

It is known, from the types of analyses illustrated earlier, that customer evaluation of the repair experience is heavily influenced by the total duration of the repair, from the time of its reporting until its official completion, or clearing time. However, it is not known whether the company standards set for repair time are acceptable to customers, nor is it known how quickly customer ratings deteriorate in the face of increasing repair times.

The company would like to know if this ratings deterioration is constant for all types of repairs. The graph below shows the relationship between reported duration and customer rating for each of several symptoms initially reported by the customer.

Figure 3
Rating Change for Different Troubles



Note that when the problem is a dead line, ratings for the repair drop drastically after only about 10 hours. On the other hand, when there is some problem with the telephone equipment, ratings are roughly stable for about 28 hours. Noise and ringing problems (where the phone does not ring for an incoming call) are in intermediate positions. The temporary rise in ratings for out-of-service problems from 8 hours to 18 hours is probably due to the necessity of the repair center's making a preliminary diagnosis before 8 hours, whose treatment is concluded before 18 hours; similar repairs taking only 8 hours generally have no such earlier communication. All ratings, of course, show a general drop-off as time increases.

THE ROLE OF CUSTOMER SATISFACTION IN SETTING PROCESS SPECIFICATIONS

Sometimes the absence of expected links between internal and customer survey measures leads to valuable insights into the way customers react to grades of service. In early 1992, a service guarantee program was introduced for small business customers in a small area in the American northwest. Consequently, several databases were created which contained various kinds of customer information. For each small business customer requesting a repair during that period, internal repair data were extracted from an internal company database. This dataset contained such data as reporting, disposition and clearing times, type (including out-of-service status) of trouble reported, and information on the final disposition and cause of the reported trouble.

One important specification for telephone repair is the creation and achievement of a verbal commitment to a particular repair time. Of these two, meeting the stated commitment is paramount, and it is highly interesting to relate the customer's and the company's measurement of how well this is achieved. One might expect that the customer's rating of the company's meeting repair deadlines should be related to the internal record of commitments met or the interval by which the commitment was missed. Surprisingly, there is no systematic relationship between these two variables.

There is, however, a clear association between repair duration and DEADLINE: "How would you rate . . . on meeting repair deadlines?"

Clearing Time					
4 hrs. or less	4-8 hrs.	8-16 hrs.	16-24 hrs.	24-48 hrs.	48 hrs. or more
Excellent					
50.3%	41.4%	33.3%	44.2%	40%*	20%*
Good/Excellent					
84.9	96.5%	77.8%	76.8%	60%*	80%*

*: based on <5 subjects

The large decrease in G/E rating as repair durations exceed 8 hours (a working day, perhaps) suggest that the repair deadlines business customers consider are those imposed by their own businesses, and not that of the telephone company. It follows that, as far as quality service is concerned, it is much less important to unilaterally impose and meet a commitment than it is to complete the repair as quickly as possible. The effect of long repair times on customers will be examined in a more dramatic way in the next section.

RELATING OPERATIONS TO OUTCOMES: INVOKING A REPAIR GUARANTEE

Based on customer survey and billing results, on anecdotal evidence, and on widespread customer confusion and dismay over repair bills complicated by telephone deregulation, a repair service guarantee was piloted for the local telephone franchise furnishing the internal/external data analyzed in the previous section. Essentially, a repair to a small-business phone was unconditionally guaranteed in the sense that repair charges were waived for any customer so requesting.

Those customers invoking the guarantee during the first five months of 1992 were matched to data from the customer satisfaction survey each completed as a repair customer, and matched again with the internal repair data described above. Because the customer base and the treatment period are both clearly defined, the variables associated with the guarantee invocation can be determined by analysis as a cohort-type study. See Hosmer and Lemeshow (1989). As always, the

reader should be aware of the general assumptions involved in these analyses: an underlying multinomial distribution, and appropriateness of the logit link function are the two most obviously technical requirements. For convenience the effects of the repair attributes are presented in tabular form with categorized data, while the reported tests of statistical significance are based on logistic regression models.

The initial status of telephone service, that is, whether the line is out of service (OOS) or not out of service (NOS) significantly affects whether the repair guarantee is invoked. The relative invocation percentages are shown in the following table, arbitrarily setting the NOS percentage to 1.0:

Service Status	Invocation Percentage
NOS	1.0
OOS	2.56

so that the probability of invocation is 2.56 times more likely in an OOS situation than NOS.

Obviously, a problem resulting in a service outage is more likely to generate a claim for the repair guarantee than some other problem, perhaps because an outage represents a potential loss of business for the consumer. If the latter hypothesis were true, then one might also expect an effect due to the duration of the outage. This is explored next.

The duration between the time of the trouble report and its clearing is associated with the likelihood of invocation, and the effect is different depending on the customer's service status. The following table gives invocation percentages for repair durations for the two service statuses, here setting the "Clearing Time $\leq$4 hrs" for OOS status to 1.0:

Service Status	Clearing Time: 4 hrs. or less	4-8 hrs.	8-16 hrs.	16-24 hrs.	24 hrs. or more
NOS	1.08	0	0	0	2.25
OOS	1.00	4.84	3.90	2.95	18.3

so that the invocation likelihood for those with OOS service cleared in more than 24 hours is 18.3 times greater than OOS status cleared in 4 hours or less.

For NOS problems, the invocation percentage is high for times above 24 hrs., but this is based on only one invocation, so there is little evidence that clearing time plays any part for these problems. For OOS problems, however, the likelihood of a guarantee invocation increases as the clearing time increases, abruptly jumping when the time exceeds 24 hours. This effect is statistically significant, and apparently important. Business customers are relatively likely to use the service guarantee as a recourse when their telephones are out of service for 24 hours or more.

It follows that quick repairs, especially for lines which are out of service, are vital in at least two distinct senses. The preceding section showed that customer satisfaction is a function of repair duration, with satisfaction levels dropping quickly if the repair takes more than 24 hours. In the analysis of this section, we saw that it makes immediate financial sense for the company to complete repairs within 24 hours: the guarantee is much less likely to be invoked if the repair is completed that quickly.

As a final point, we note that there is evidence that customer satisfaction ratings have an important role in the link between service improvement pilots and customer behavior. In this situation, there is no effect of the institution of the guarantee program on customer satisfaction, either generally, or specifically for repair service. Satisfaction with the repair service is, however, strongly related both to invocation of the guarantee and to the internally measured characteristics which we showed were associated with invocation.

Since most of the data from the survey, the internal repair characteristics, and the invocation behavior are categorical, and not easily transformable into a normal distribution, structural equation modeling is not appropriate to study the relationships among these three sets of data. Logistic regression, however, is available to model the individual links. The table below shows the strength of the links between invocation and CS, and between invocation and the internal characteristics of repair duration and initial status (OOS/NOS). Strength is measured in terms of the size of -2 log- likelihood (which increases with the goodness of the fit) and the percentage of

observed/predicted discordant pairs (which decreases with increasing goodness of fit).

Model	-2logL (degrees of freedom)	% Discordant Pairs
Invocation = f(OOS/NOS, Duration)	9.908 (2 df)	16.8
Invocation = g(CS:%G/E)	74.529 (1 df)	0.8

The good fit of the Invocation-CS model, and the rather more mediocre fit between invocation and the internal measures strongly suggests that CS is a mediating influence between the objective characteristics of this pilot and customer behavior. The characteristics affect customer satisfaction, and this attitude in turn affects the invocation behavior.

CONCLUSIONS

We have described and illustrated each major step of the feedback loop which links customer satisfaction data with internal company operations, both routine and pilot programs, and then back again to customer evaluation and behavior. In this way, customer satisfaction data fulfills its central role in modern quality improvement theory, and becomes extensively useful to the company.

In using customer data to improve operations, as TQM philosophy prescribes, customer satisfaction measures should be related to customer perceptions of operations to provide a clear notion of the relative importance of service attributes in the customer's mind. When they exist at the appropriate level, internal company measures can usefully be linked to satisfaction data to provide a firmer and more objective target for company improvements in operations. The examples analyzed here quantified the importance of repair times for several types of telephone problems. As satisfaction was related back to repair times, those times can be related forward beyond satisfaction to customer intention and behavior. Our pilot program of instituting repair guarantees gave significant indications of the role of internal repair characteristics to both satisfaction and guarantee invocation, thus giving the company firm information on the revenue implications of this program.

REFERENCES

Anderson, E. and M. Sullivan (1993), "The Antecedents and Consequences of Customer Satisfaction for Firms," *Marketing Science*, 30, 125-143.

Bearden, W. and J. Teel (1983), "Selected Determinants of Customer Satisfaction and Complaint Reports," *Journal of Marketing Research*, 20, 21-8.

Bolton, R. and J. Drew (1992), "Mitigating the Effect of Service Encounters," *Marketing Letters*, 3, (1), 57-70.

Deming, W. Edwards (1986), *Out of the Crisis*, Cambridge, MA: MIT Center for Advanced Engineering Study.

Fornell, C. (1992), "The Customer Satisfaction Barometer," *Journal of Marketing*, (January), 56, 6-21.

Fuller, W. (1987), *Measurement Error Models*, New York: John Wiley and Sons.

Hauser, J. and D. Clausing (1988), "The House of Quality," *Harvard Business Review*, (May- June), 63-73.

Holbrook, M. (1981), "Integrating Compositional and Decompositional Analyses to Represent the Intervening Role of Perceptions in Evaluative Judgments," *Journal of Marketing Research*, 18, 13-28.

Hosmer, D. and S. Lemeshow (1989), *Applied Logistic Regression*, New York: John Wiley and Sons.

LaBarbera, P. and D. Mazursky (1983), "A Longitudinal Assessment of Customer Satisfaction/Dissatisfaction: The Dynamic Aspect of the Cognitive Process," *Journal of Marketing Research*, 20, 393-404.

Narasimhan, C. and S. Sen (1990), "Linking Engineering Attributes to Perceptual Characteristics," Working Paper, Washington University at St. Louis.

Neslin, S. (1983), "Designing New Outpatient Health Services: Linking Service Features to Subjective Customer Perceptions," *Journal of Health Care Marketing*, 3, (Summer), 8-21.

Parasuraman, A., Valerie Zeithaml and Leonard L. Berry (1986), "SERVQUAL: A Multiple-Item Scale for Measuring Customer Perceptions of Service Quality," working paper, *Marketing Science Institute*, Research Program, Cambridge, MA, Report #86-108, (August, 1986).

Walton, M. (1986), *The Deming Management Method*, New York: Perigee Books.

Send correspondence regarding this article to:
James H. Drew
GTE Laboratories Inc.
40 Sylvan Road
Waltham, MA 02254