CUSTOMER SATISFACTION AND THE PREDICTION OF BEHAVIOR IN A HYPOTHETICAL CONTEXT

James H. Drew, GTE Laboratories Incorporated
Tina M. Bronkhorst, GTE Laboratories Incorporated

ABSTRACT

From a company perspective, customer satisfaction and the subjective judgment of general quality are most useful when viewed as part of a chain of customer evaluations ranging from attribute perceptions up to loyalty and future behavior predictions. While they are of paramount interest to the manager, the latter items are notoriously difficult to measure well, so that their intrinsic meaning is unclear and their linkages to more familiar constructs are tenuous. These problems are vastly inflated when the behavior of interest is necessarily hypothetical. This paper describes several measures of intention in a hypothetical context, exhibits ways to ensure and check their validity, and discusses their relationship to lower-level customer judgments.

INTRODUCTION

Customer satisfaction is, by definition, a post-purchase affect, and its measurement is naturally reactive. This makes its measurement fairly easy for the customer and often its interpretation for the provider straightforward, at least when the structure of its current service is at issue. However, most managers advocate its use as an indicator of such future behavior as service usage or loyalty. Thus, for the practitioner, a crucial issue in the study of satisfaction and service quality is the link between these concepts and those of future behavior.

The difficulty, both for the customer and for the manager, of translating a retrospective attitude into a future state is enhanced when the service environment is undergoing fundamental changes, so that the customer must not only predict his future behavior in a stable business environment, but must also account for radical changes in the source, bundlings and form of the service. The question must be valid in the two senses of 1) measuring the proper customer understanding of the prediction task, and 2) predicting behavior under that environment.

Telecommunications is undergoing a revolution that is unique in creating unconventional services from unconventional providers, yet in many ways appears to look like the plain old telephone service with which consumers have been familiar for many decades. The challenge for the satisfaction researcher is making sense of the highly hypothetical questions put to the industry's current customers, for overcoming this issue is the key to going beyond satisfaction and developing measures that more directly affect the viability of the provider's organization.

In this type of environment, customers are asked to evaluate many dimensions of the telecommunications industry: functional attributes which examine how the service performs, interaction or transaction variables, which examine the perceptions of any interactions the customer may have experienced with the service provider, behavioral intentions , which attempt to measure future behavior, and image questions, which ask about impressions of the corporation. The last two types of questions tend to be phrased hypothetically (e.g., "How likely would you be to recommend GTE to a friend or business associate?")

The research reported here utilizes data from two studies of telecommunications service (for convenience, labeled Study I and Study II) collected during 1994. Study I addressed issues of customer satisfaction with GTE's core service attributes, as well as items typically labeled as "image" items (e.g., GTE's leadership in the community.) The survey was administered to 384 GTE residential customers. Study II, administered to 1171 residential and business customers of GTE, focused on the comparison of GTE with other providers of telecommunications services, namely long distance providers, cellular service providers, and cable television providers.

The paper is organized in the following manner. After briefly discussing the relevant literature in this domain, we will examine how hypothetical questions can pose response problems for respondents and demonstrate how these problems may be overcome with appropriate question positioning and structure. Next we will

discuss the validity of the data examined in this manner. Models of loyalty for local and long distance telephone service will be presented, and the relationship of these concepts to the more popular items of service quality and value will be examined. In the three loyalty concepts the differences between an implicit and an explicit consideration of alternative suppliers will be explored. This analysis will be seen to have implications for the revelation of customer behavioral intentions, for the place of service quality in those intentions, and for the bases on which loyalty decisions will be made.

PREVIOUS WORK

Predicting the behavior of a customer based on his or her current perceptions and preferences is generally not easy, and relationships between affects, cognitions and behaviors are hard to quantify. Morwitz and Schmittlein (1992) point out that most practical methods of converting buying intentions into actual purchase rely on past purchase behavior or unreasonably arbitrary assumptions about the customer's affect intensity and purchase. When the behavior in question is not currently possible, such as the purchase of a new product, its prediction is more difficult for the customer and the interpretation of his/her prediction is problematic for the researcher. Great pains are taken, for example, to accurately estimate "intent to buy" reactions from customers, and the physical simulation of such hypothetical purchases (Silk and Urban, 1978) has been shown to be especially effective. However, we consider a situation where such simulation is not possible at all. Not only is the proposed service not yet available, but its form and the business environment in which it might be delivered are under construction. Of necessity, one must rely on verbal descriptions followed by the questioning of potential customer behavior. In that regard, past research has also shown the relationship between loyalty and image (or brand reputation) (Selnes, 1993).

There are several behaviors of particular interest to an organization whose industry is undergoing redefinition. In such times, the classical concepts of voice (e.g. word-of-mouth complaining, extolling, recommending) and exit (unsubscribing, selective by-passing) as described by Hirschman (1970) are supplemented by the possibility of cross-industry switching, where a current service is provided by an unconventional supplier, or the conventional supplier provides service hitherto supplier elsewhere. As an example of the latter behavior, post office package delivery in the 1970s became replaceable by UPS and Federal Express, while the US Post Office later added express mail delivery.

For a conventional provider, the anticipation and preparation for these three customer behaviors is a crucial project. Their measurement takes on a particular hierarchy for a regulated utility, where customer exit is impossible or extraordinarily difficult. As shown by Fornell and Bookstein (1982), voice predominates when exit is blocked, and a popular measure of customer loyalty--or potential loyalty--is whether the customer would recommend the service in question to a friend who did have a choice of providers.

THE INCIDENCE OF DIFFICULT MEASURES

Some concepts can be very difficult for customers to understand and respond to, even for an offering as familiar as telephone service. A prime indicator of a problem question is the number of missing values it yields. A study of telecommunications services conducted by the authors (which, for convenience, will be called Study I), asked GTE customers to evaluate many aspects of their local telecommunications provider, including items about company image and behavioral intentions. Table 1 includes a selection of those survey items, and their paraphrased wordings, and shows the number of missing values.

Although the proportions of missing values for RECOMMEND and CHOOSE are at a reasonably low level, the items requiring the consideration of other companies, such as BEST CORPORATION, the explicit naming of other companies (COMP. REPUTATION), and the evaluation of other companies (ONE_SUPPLIER) have unacceptably high levels.

Apparent reasons for high numbers of missing values are:

Table 1
Study I - Missing Values

Question	No. of Missing Values (Base of 384 unless noted.
CHOOSE If you were able to choose your own telephone company, how likely would you be to continue using GTE?	9.1% 35/384
RECOMMEND If the telecommunication industry changed so that other companies besides GTE also provided telecommunication services in your area, what is the likelihood that you would recommend GTE to a friend or someone else?	6.5% 25/384
BEST CORPORATION Next I would like you to think about any large national corporation - companies of any type of business that you might think of. When you think of large national corporations which one corporation do you feel has the best overall reputation?	40.6% 156/384
COMP. REPUTATION Compared with [...], how would you rate GTE's overall reputation?	24.1% 55/228
ONE_SUPPLIER (Would you choose GTE as a single supplier of local, long distance, cellular and CATV)	60.4% 232/384

Hypothetical situations This would be relevant in the responses to the questions CHOOSE, RECOM, and ONE_SUPPLIER. In response to these questions, customers must envision a business environment currently not in existence, and whose future dimensions are unclear (or perhaps even unknown.)

Requires Specialized Perspective The responses to the questions COMP. REPUTATION demands that the customer have some knowledge of the corporations function, for example, within a community or among others in the industry.

Requires too much cognitive processing The questions BEST CORPORATION and COMP. REPUTATION force a customer to consider 1. other corporations in any industry, or 2. their previous supplier of telecommunications services. The activity of recalling this information and evaluating its relevance (within the scope of a telephone interview) could prove too taxing a process.

Note that items relating to competition and alternative providers constitute a sizable portion of these problems. Yet it is precisely these issues that are of most interest to the organization in evaluating its long-term viability, for these affect customer behavior in a competitive environment.

Such reluctance to respond to the alternative supplier items, however, is not inevitable. In the above survey, most of the items were concerned with GTE's performance, unrelated to any reference points. In responding to those few items where comparisons were important (BEST CORPORATION, COMP. REPUTATION), the customer was not prepared to consider alternative suppliers and integrated service.

In contrast, a second residential survey (Study II) began by asking customers for names of their

long distance, cellular and CATV providers. Customers were then asked for ratings of each cited provider, *prior to* items associated with the integration of their services and customer choices among the consequent competitors. The questionnaire for this study employed the following devices to enhance validity of the hypothetical behavior items:

1. calling to mind who the other service providers in a market may be,

2. defining the customer's perspective,

3. providing cues to aid in the recall and evaluation of relevant information called to mind, and

4. positioning items to promote attribute comparisons among the various telecommunications providers.

As seen in Table 2, this positioning of the questions addressed our three concerns above:

Table 2
Study II - Missing Values

Item	Percent Missing
CHOOSE	8.0% (93/1171)
RECOMMEND	13.7% (160/1171)
SINGLE- (If you could obtain your local , LD, Cellular and CATV from a single supplier, would you do so)	7.7% (90/1171)
ONE_SUPPLIER- (If you could choose one supplier, who would you choose)	29.3% (257/877)

The proportion of missing values for the current provider-focused items RECOMMEND and CHOOSE are about the same as for Study I, as one might expect, but substantial improvements are apparent in the proportion of missing values for SINGLE and ONE_SUPPLIER, where other suppliers must be explicitly considered.

This is the first indication that these hypothetical behaviors can be answered by residential customers. Cognitive processing demands can be made sufficiently manageable so that answers can be supplied. The next issue is whether these answers are valid in the sense that they address the intent of the question.

THE VALIDITY OF TELEPHONE CUSTOMER HYPOTHETICAL INTENTIONS

The validity of responses to hypothetical questions about future behavior is not often discussed in many marketing and social psychology experiments. The issues here, comprising the production of a thoughtful, coherent response to a future behavior question as well as the predictive validation issue of whether the loyalty intention will be realized, are particularly prominent for local telephone service. Not only are switching possibilities highly novel, but the passivity of the service does not promote profound consumer thought.

In the past, several approximations to loyalty behavior have been created to overcome the extreme hypothetical nature of brand switching from a regulated monopoly. Customers were asked whether they would recommend the local telephone service to a friend who had a choice, while other customers were asked to imagine they were able to choose among current local exchange carriers. (Note that both are rough approximations to the desired information. In the first, the act of switching is transferred to another person, and in the second, only local exchange carriers (LECs) are allowed as switching destinations.)

Recommending GTE's service to a friend with a choice, or choosing among different local exchange carriers is mostly a hypothetical situation, so it is not possible to directly verify that a customer would carry through an intent to recommend or choose GTE, and we must therefore rely on indirect suggestions. First, measures must be correlated with each other, for they estimate similar concepts. Table 3 shows the correlation

between Recommend and Choose for each of the four alternative providers surveyed, for the residential market (recall that a correlation of one signifies perfect linear agreement, and a correlation of zero connotes complete independence):

Table 3
Correlations Among Measures

Service	Correlation between Recommend, Choose
GTE	0.675
Long Distance	0.639
Cellular	0.713
CATV	0.777

Recommend and Choose are as closely related for local service, for which there is no choice, as for Long Distance, for which there is. This suggests that both items are measuring nearly the same concept, and since this concept is actual choice or recommendation for LD and Cellular, the concept is almost surely intended choice.

Second, items measuring different concepts should be uncorrelated, indicating that the survey or customer defaults are not driving the Recommend and Choice responses. In fact, the four Choice variables and the four Recommend variables for each provider are not highly correlated, the general level being only 0.200 to 0.400. In other words, similar survey items which should address different customer thoughts, are in fact not highly related.

Finally, the regression models for Recommend and Choice for the two suppliers for which choice is a possibility should not be appreciably better (in terms of R^2) than the models for local service. The R^2 values, which in this case are upper bounds for measures of fit for each of the four services, are as follows:

Table 4
R^2 Coefficients for Choice and Recommend

Service	Maximum R^2 Residential
Choice-GTE	0.407
Choice-Long Distance	0.300
Choice-Cellular	0.388
Choice-CATV	0.542
Recommend-GTE	0.438
Recommend-Long Distance	0.366
Recommend-Cellular	0.420
Recommend-CATV	0.643

Note that the Recommend and Choose models for GTE fit at least as well as those for the LD and Cellular models, where there is a real choice (i.e., not a hypothetical choice). Models for Small Business and Medium/Large Business customers show almost exactly the same pattern. This is further evidence that Recommend and Choose have at least as much meaning for GTE-based responses as for other suppliers.

THE ANTECEDENTS OF CONSUMER INTENTIONS

The previous sections demonstrated that intentions in hypothetical situations can indeed be validly measured, using the proper procedures. Now we discuss the antecedents of the intentions measures introduced above.

We consider whether quality and value constitute a summaries, or mediator positions between the attributes and the intention items. Alternatively, the intention items may be at least partially directly influenced by the attributes

themselves. This is parallel to the discussion of mediating variables in Baron and Kenny (1986).

Consider models of Choose and Recommend whose explanatory variables include all (potential) lower-level survey variables. Thus, Quality and Value are included in each model, as are the attribute variables Cust. First, Technical Leadership and Reliability. These analyses yield the following coefficients (asymptotic t-values shown in parentheses, where the R^2 values are derived from the individual OLS regressions):

Note that the hypothetical behavior variables apparently depend directly on some of the perceived attributes, and Price (meaning the reasonableness of the price charged) depends on those same attributes as well as on Quality. The "intermediate" concepts of Quality, Value and Price evidently capture aspects of the service's performance not included in the explicit attributes. However, it is important to notice that in the determination of RECOMM and CHOOSE, Value is not the most important input, as has been postulated (See, for example, Holbrook, 1994). On the other hand, these same intermediate concepts do not constitute a cognitive summarization of the services suitability to enter into the behavior evaluations in isolation, since the original attributes also enter into the hypothetical behavior questions. To make the customer-perceived service attributes useful, they must be related to the functional aspects of the service, preferably categorized along organizational lines.

Note also an interesting set of differences between the Choose and Recommend construct: the latter puts greater weight on reliability, quality and price than the former, and after the effect of Recommend is accounted for in Choose, the subjective variables of Value and Puts Customer First are the only ones still significant. Apparently, recommendation requires basic adequacy of the service here, but personal choice additionally requires high customer orientation and a favorable personal utility/personal cost ratio.

THE EFFECT OF ALTERNATIVE PROVIDERS

The previous section discussed the position of the concept of service quality in the structure of items concerned exclusively with a single company, namely the customer's current provider for each of several aspects of telecommunications service. Now, consider the role of questions regarding other, potentially alternative suppliers.

We have focused on measures of the hypothetical behavior of customer loyalty in terms relating to the current provider (i.e. the local telephone company) only. However, the data we have analyzed above contains three distinct measures of customer behavioral intention toward GTE, which have traditionally been termed measures of loyalty. As previously described above, the first two, Recommend and Choose, are direct items in the survey.

Table 5
Model Coefficients - Local Provider (GTE)

	Explanatory Variables							
Dependent Variable	R^2	Recomm	Value	Price	Quality	Cust. First	Tech. Lead.	Rel.
Choose	0.505	0.423 (13.06)	0.098 (2.53)	0.044 (1.30)	0.065 (1.77)	0.127 (3.40)	0.059 (1.70)	0.043 (1.34)
Recomm.	0.437		0.128 (3.11)	0.162 (4.56)	0.214 (5.55)	0.142 (3.61)	0.040 (1.08)	0.129 (3.79)

The third measure is a composite of two items. A customer is considered loyal if he/she answers "Extremely . . .", "Very . . ." or "Somewhat Likely" to the item:

> "If you could obtain your local telephone, long-distance, cellular or cable TV services from one supplier, what is the likelihood that you would do so?"

and subsequently identifies that single supplier as GTE. In the analysis below, this two-valued measure is named SINGLE.

From the wording of the items, it appears that CHOOSE and SINGLE differ in that CHOOSE offers the customer a choice of current local telephone providers, while SINGLE widens the choice to current suppliers of LD, Cellular and Cable providers.

The three measures of loyalty (Recommend, Choose and Single Supplier) are functions of very different kinds of variables. While Choose and Recommend depend only on GTE's own ratings, as detailed above, the choice of a single supplier depends on the comparative scores given quality or customer orientation ratings. Comparative scores, coded BEST_XXX, are exemplified by:

> BEST_QUAL: GTE had the best (i.e. strictly better) QUALITY score among the local telephone, long distance, cellular, cable providers rated, and
>
> BEST_PRICE, defined similarly.

The SINGLE variables is found to have the following model structure, based on a logistic regression.

In contrast to the RECOMMEND and CHOOSE measures, two of the three inputs are *comparative* ratings: for GTE to be chosen as the single supplier of local, LD, cellular and cable services, its quality must be strictly higher, and its prices strictly better than the competition. In addition, its technological leadership must be high, though not in a comparative sense, presumably to unite these different services in one package.

This analysis, which yields similar results for customer choice of other potential suppliers, is important not only for its indication of service attributes used in this hypothetical choice activity, but for its strong suggestion that such choice is dependent on service beyond that measured by one company's quality ratings.

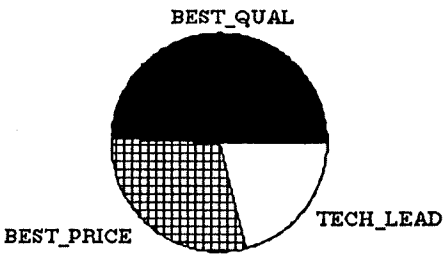
CONCLUSIONS

We conclude that:

> Telephone customers can meaningfully respond to hypothetical questions about future behavior, when proper survey procedures are followed.
>
> Service quality is not consistently an important ingredient in determining customer's intended behavior. More specific service attributes are of comparable importance, and suggest that

Table 6
SINGLE-Residential Customers

SINGLE	ATTRIBUTE	COEFFICIENT
	BEST_QUAL	1.049
	BEST_PRICE	0.616
	TECH_LEAD	0.448



this concept is useful as a supplement to a necessarily limited attribute list, rather than as an intermediate summarization feeding into higher level decisions.

In particular, increasing the scope of one's service to include the those currently offered by other types of providers depends on the customer's *comparative* satisfaction and his/her *comparative* rating of certain key attributes.

It follows that from a business perspective, the measurement of customer satisfaction is a fundamental requirement for favorable future behavior, but is by no means sufficient. Even though retrospective satisfaction and service quality are relatively easy to make meaningful and measurable for the customer, understanding the future health of the company is dependent on customer-oriented concepts that are vague, hypothetical, hard to interpret--and crucial.

REFERENCES

Baron, R. M. and D. A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic and Statistical Considerations," *Journal of Personality & Social Psychology*, 51 (6), 1173-82.

Fornell, C. and F. L. Bookstein (1982), "Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory," *Journal of Marketing Research*, XIX (November 1982), 440-52.

Hirschman, A. O. (1970), *Exit, Voice and Loyalty: Responses to Declines in Firms, Organizations and States*, Cambridge, MA: Harvard University Press.

Holbrook, M. (1994), "The Nature of Customer Value,": in *Service Quality:New Directions in Theory and Practice*, eds. R. Rust and R. Oliver, Thousand Oaks, CA: Sage Publications.

Morwitz, V. and D. Schmittlein (1992), "Using Segmentation to Improve Sales Forecasts Based on Purchase Intent: Which "Intenders" Actually Buy?," *Journal of Marketing Research*, XXIX (November 1992), 391-405.

Selnes, F. (1993), "An Examination of the Effect of Product Performance on Brand Reputation, Satisfaction and Loyalty," *European Journal of Marketing*, 27, 19-35.

Silk, A. and G. Urban (1978), "Pre-Test Market Evaluation of New Product Goods: A Model and Measurement Methodology," *Journal of Marketing Research*, XV (May 1978), 171-91.

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