CHAPTER 10. CONVERTING SERVICE QUALITY RESEARCH FINDINGS INTO TRANSIT AGENCY PERFORMANCE MEASURES

10A. Introduction

The assessment of the determinants of transit service quality has so far focused on the analysis of the survey that measures transit users' attitudes towards service and derives the implied priorities for transit service improvements. This analysis has provided useful insights into the factors that make up transit rider satisfaction which influence mode choice behavior and consequently the observed transit ridership.

The interpretation of the survey results by managerial staff in each of the three transit agencies further underscores the importance and usefulness of developing and maintaining a market research program that focuses on customer satisfaction. The robustness and resonance of the survey findings with management's opinions about the service offered bring to focus the steps that are required to take action to improve service.

In this chapter we build upon the existing analysis framework by structuring the discussion of performance measurement from a transit agency's management perspective. Instead of focusing on the quality of service perceived and expected by the customer, we shift to ways of measuring the quality of service actually offered by the transit agency. The ability to accurately measure performance allows the agency both to evaluate its service and to define realistic and measurable goals for service improvements.

We first discuss the importance of linking transit riders' perspectives to objective disaggregate measures of transit performance. The different types of analyses that can be conducted are discussed along with the desired elements of an ongoing data collection plan that focuses on the greatest possible level of detail.

The performance measures are then identified in a manner that is consistent with customers' experience by breaking down a transit trip to its individual components and by defining customer expectations of service. Each of the 46 transit service attributes that were evaluated in the survey is related to the different components of the transit trip to identify service attributes that share common characteristics.

The 10 most important aspects of service that have been identified through the survey analysis for each transit agency are then tabulated to identify service attributes that are common to rail and bus transit systems in each of the three cities. For each of those service attributes we define customers' expectations and discuss a range of mostly simple performance measures that can be used to measure the ability of the transit agency to offer service that meets these expectations.

10B. A Transit Agency's Perspective to Transit Performance Measurement

The consumer-oriented approach to transportation service planning is rooted in the assumption that the observed transit ridership and transit market share are the result of the mode choices made by each individual commuter. The analysis framework presented in Figure G.1 of Appendix G highlights the importance of transit level of service, individual traveler characteristics, and communication and marketing channels on the formation of travelers' perceptions and consequently on their likelihood of riding transit.

The analysis of the transit rider survey has provided a way of evaluating the link between riders' perceptions and their overall satisfaction with transit service. A better understanding of transit customers' needs and wants would allow the transit agency to identify the strengths and weaknesses of transit service against competing modes and the differences in service for individual routes within the transit system.

Examples of successful customer-driven approaches to the design and marketing of transit service quality are documented in a recent study of four European transit systems.[14] The common theme among these case studies is the intent to demonstrate the transit agency's commitment to service quality and its sensitivity to customer input by promising a standard of service. This allows customers to evaluate the ability of the transit agency to provide the level of service to which it was committed.

Among the service improvements that were considered and implemented in the transit systems under study were the provision of more frequent service, the improvement of reliability, purchase of new equipment, improved customer relations, electronic payment facilities, and more convenient connections. A similar review of 40 transit systems in the United States[15] identified increases in transit ridership that system managers attributed to level of service adjustments, pricing changes, marketing and information initiatives, enhancement of service coordination, and market segmentation.

Therefore, the next important step in the process from a transit agency perspective is to develop a strategy of service improvements that is responsive to its customers' expressed needs and wants. In particular, a transit agency needs to define the type and level of service improvements that need to be implemented to address weaknesses in service for those service attributes considered most important by its customers.

The collection of data reflecting riders' perceptions of transit service along with an ongoing program of transit performance data collection at the transit line and route level by different times of day and days of the week can be used by a transit agency to:

- identify existing weaknesses of transit service as reflected in the responses provided by transit riders and in the performance measures being monitored;

- set priorities for service improvements by focusing on the aspects of transit service that need to be addressed first and by identifying the service routes and segments of the market that will be affected the most;

- design and implement the identified improvements in transit service; and

- design an information dissemination program that will properly communicate the improvements to the riding public.

A recent Transit Cooperative Research Program study[16] approaches the subject of quality of transit service by adopting a total quality management (TQM) framework for public transportation. To meet the objectives of increased productivity, reduced costs, and higher ridership through improved rider satisfaction the study focuses on controllable factors that influence public transit performance. Recognizing the human service character of public transit, the study focuses on "putting customers first" by responding to customer expectations and by translating market research into actionable procedures.

An important consideration in the outlined approach is the ability to "manage by fact" and establish a range of measures that can be used to monitor and evaluate performance. Among the criteria for developing these performance measures that are included in the report are the:

- validity of data that are sampled by credible unbiased methods;
- completeness of data that cover a broad spectrum of aspects of service;
- policy sensitivity of data that can be used to support managerial decisions;
- timeliness of data that can be processed, analyzed and interpreted on time;
- transparency of the data collection process;
- inexpensive data that may already be collected for another purpose; and
- ability to interpret data by developing measures that are easy to understand, compare, and communicate to management and the public.

The ability to make the linkage between riders' statements and measures of transit performance is therefore instrumental in providing transit management with the means of evaluating alternative service improvements aimed at enhancing rider satisfaction and transit ridership. Such an evaluation can be supported by an ongoing data collection effort that captures differences by transit route, time of day, and day of the week and focuses on a comprehensive list of transit performance indicators. As a result, the ongoing analysis of the transit performance measures can be used to:

- provide transit management with a systemwide overview of transit operations for different transit modes;
- evaluate transit performance on a route-specific level of detail by focusing on individual segments of the transit network;
- monitor changes in transit service over time to identify deteriorating conditions or to highlight improvements in service in response to service intervention;
- identify the variation in transit level of transit service by collecting data specific to a service area, time of day, or day of the week for the service attributes of interest; and
- guide the development of marketing and communication strategies to inform transit customers and potential customers of the desirable service features.

10C. Overview of Transit Performance Measures

The collection of transit performance data to support the monitoring, evaluation, and the implementation of improvements in service presents a challenge to transit agencies. Although transit agencies might be interested in collecting a wide array of information, the cost of collecting and analyzing a large amount of transit performance and service quality data presents a constraint to transit agencies.

As a result, the data collection and analysis activities should be concentrated on those aspects of transit service that are both crucial to their operations and that more accurately reflect the needs and wants of customers and potential customers. The objective is to match the most important perceptions to specific aspects of transit service and to identify one or more corresponding service performance indicators. These measures will differ by transit agency given the different priorities expressed by riders, the differences in the nature of services offered, and the feasibility and cost of collecting the relevant data.

Travelers' need to travel reflects their need to participate in an activity that is located elsewhere. In this context, travelers' choices of residential location, workplace, time-of-day of travel, and transportation mode reflect their desire to minimize the disutility of travel. In the case of transit riders, the disutility of travel encompasses the whole travel experience from the planning of a transit trip at their point of origin through the walk egress portion of the trip to get to their final destination. To better understand and measure the service that a transit rider receives, the total travel experience has been broken into the trip components and service dimensions shown in Table 10.1.

Table 10.1
Correspondence Between Trip Components and Dimensions of Service

Trip Components	Service Dimensions
Trip planning	Passenger information
Cost of transit	Fare level and type
Walk to transit stop	Accessibility
Wait at transit stop	Station environment Passenger information Service delivery Security
Travel by transit	Vehicle environment Passenger information Service delivery Security
Transfer to transit	Station environment Passenger information Service delivery Security
Walk to destination	Accessibility

Prior to their trip, transit riders may need to seek information about the most convenient route, departure time, transfers, and fare to get to his or her destination. Sources for such information include printed transit route maps and schedules, information provided over the phone by individuals at a passenger information center, and electronic versions of schedule and fare information. Although such information is seldom needed for routine trips, it can be of great value to infrequent transit users and non-users who are unfamiliar with the system.

The level of transit fares is another aspect of transit service that contributes to the disutility of travel and affects riders' perceptions of transit's attractiveness. Although transit fares are often lower than the corresponding operating, maintenance, and parking costs of private modes, fare levels can have an adverse impact on the price-sensitive frequent traveler segment of the travel market. The availability of different types of fares, such as monthly passes, ten-ride discount tickets, and electronic fare media with value storage capabilities, and fare restrictions increase travelers' flexibility to choose an optimal payment strategy that fits their own travel patterns.

The travel components of a transit trip include:

- the access to the transit station/bus stop,
- the time spent waiting for transit service,
- the in-vehicle experience of riding transit,
- potential transfer(s) to different transit services, and
- the egress to the final destination.

The access and egress walk components of the trip are only in part linked to the everyday operations of a transit system. Although the number, location, and spacing of stations and stops and the adjacent land-use development may affect transit service considerably, they are primarily linked to the original design of the service. On the other hand, riders' perceptions of the accessibility of rail stations and bus stops can be positively influenced by interventions such as kiss-and-ride facilities, availability of long-term station parking, sidewalk maintenance, availability of well-lit access paths, and maintenance programs for stairs, escalators, and elevators leading to platforms.

The time waiting at the station or stop, the in-vehicle component of the trip, and the transfer to another transit route are all characterized by:

- traditional measures of transit service such as wait time, travel time, and service reliability;
- the station/stop and vehicle environments that the transit riders experience; and
- the availability and quality of information available to riders at rail stations, bus stops, and en route.

Table 10.2 provides a link between the components of a transit trip, the dimensions of transit service, and the 46 attributes of service that were used in the transit rider survey. These linkages illustrate both the depth of the rider survey and the potential range of corresponding measures of performance. The list of candidate performance measures can be extended even further considering that a variety of measures can be defined for attributes like service reliability depending on the nature of service. A range of surrogate measures may be needed to properly reflect riders' feelings of security at stations, stops, and on-board transit vehicles.

Table 10.2
Ratings of Service By Trip Component and Service Dimension

Trip Components	Service Dimensions	Ratings	Potential Measures
Trip planning	Passenger information	Availability of information by phone and mail	Response time for providing requested information. Accuracy of information provided.
Cost of transit	Fare level and type	Availability of monthly/discount passes	Available fare types.
		Fairness and consistency of fare structures	Relative cost per mile for different traveler segments.
		Cost effectiveness, affordability, and value	Cost per one-way trip.
		Cost of making transfers	Cost of transfers.
		Ease of paying the fare	Available fare types. Fare restrictions (exact fare only, surcharges, etc.).
Walk to transit stop	Accessibility	Accessibility to those with a disability	Percent of elevators in working condition.
Wait at transit stop	Service delivery	Frequent service so wait times are short	Wait time.
		Hours of service during weekdays	Hours of weekday service.
		Frequency of service on weekends	Weekend wait times. Hours of weekend service.
		Reliable trains/buses that come on schedule	Schedule adherence measures.
	Passenger information	Availability of schedules/maps at stations/stops	Availability.
		Display of customer service number	Availability.
		Transit personnel who know the system	Percent of correct answers to questions.
		Provision of signs and information in Spanish	Availability.
		Explanations and announcement of delays	Audibility of public address system.
		Posted minutes to next train/bus	Availability, accuracy of information.
	Station environment	Availability of shelters and benches	Availability.
		Physical condition of stations/stops	Quality of lighting, seating, telephones.
		Cleanliness of stations/bus stops	Presence and amount of trash.
		Absence of graffiti	Presence and amount of graffiti.
		Absence of offensive odors	Presence of such odors.
	Security	Safety from crime at stations/stops	Presence of police, transit staff, emergency phones, etc.
		Freedom from the nuisance behavior of others	"Quality-of-life" and system rules violations.
Travel by transit	Vehicle environment	Absence of graffiti	Presence and amount of graffiti.
		Absence of offensive odors	Presence of such odors.
		Availability of handrails or grab bars	Availability. Percent of passengers with access to them.
		Availability of seats on the train/bus	Number of seats. Number of unused seats.
		Cleanliness of the train/bus exterior	Presence and amount of graffiti and dirt.
		Cleanliness of the train/bus interior	Presence and amount of trash.
		Comfort of seats on the train/bus	Material and condition of seats.

Table 10.2
Ratings of Service By Trip Component and Service Dimension
(continued)

Trip Components	Service Dimensions	Ratings	Potential Measures
Travel by transit *(continued)*	Vehicle environment *(continued)*	Ease of opening doors of train/bus	Need for special efforts to get through vehicle's doors.
		Temperature on the train/bus	Temperature range within the vehicle.
		Quietness of the vehicles	Noise level without riders.
		Smoothness of the ride and stops	Acceleration/deceleration profile of vehicles. Driver training record.
		Physical condition of vehicles and infrastructure	Maintenance profile of vehicle. Breakdowns.
	Passenger information	Clear and timely announcement of stops	Audibility of public address systems. Presence of announcements.
		Route/direction visible on trains/buses	Availability, accuracy of maps, destination indicators.
		Friendly, courteous, and quick service	Demeanor of transit personnel.
		Transit personnel who know the system	Percent of correct answers to questions.
		Explanations and announcement of delays	Presence of announcements.
		Station/stop names visible from train/bus	Number and spacing of signs, lighting.
	Service delivery	Frequency of delays for breakdowns/emergencies	Mean distance between failures.
		Trains/buses that are not overcrowded	Load factors.
	Security	Freedom from the nuisance behavior of others	"Quality-of-life" and system rules violations.
		Safe and competent drivers/conductors	Driver training and accident records.
		Train/bus traveling at safe speed	Vehicle travel speeds.
Transfer to transit	Accessibility	Connecting bus service	Availability/frequency of connecting services.
		Short wait time for transfers	Transfer times.
		Accessibility to those with a disability	Percent of elevators in working condition. Presence and quality of curb cuts, ramps, tactile surfaces.
Walk to destination	Accessibility	Accessibility to those with a disability	Presence and quality of curb cuts, ramps, tactile surfaces.

In the remainder of this chapter, we focus on the 10 most important determinants of service for each of the transit systems under study. Table 10.3 summarizes the findings and highlights the similarities and differences across the three systems and the two CTA lines that were examined.

The two service attributes that emerged as the most important across all routes sampled were the frequency and reliability of transit service, both of which reflect important policy-sensitive aspects of transit service design. The third service attribute that was mentioned by riders in all three transit systems but only in one of the CTA lines was the freedom from the nuisance behaviors of others, an important but subtle and difficult to quantify service dimension. The remaining "top ten" service attributes were split between those that were perceived as important by riders in Chicago and Lynchburg and those that were mentioned by riders of the Sun Tran service who mostly focused on frequency-related issues.

In sections 10D to 10M, we focus the discussion on the individual service dimensions and the corresponding measures.

10D. Frequency of Transit Service

Based on the customer satisfaction surveys, frequency of transit service is among the most important elements of transit service. Frequency was at the top of riders' lists for each of the three agencies where transit riders were surveyed.

Frequency has two interpretations for transit riders. First, it refers to the hours of operation of transit services. Many routes and services are available only during weekday peak periods, and sometimes riders need to make trips served by the routes and services on weekends and on off-peak times of weekdays. Limitations in transit service hours obviously affect travelers who need to travel during the hours or days when there is no service. In addition, some potential transit riders choose not to use transit services because the particular services are unavailable for their anticipated return trips or because they cannot be certain about the time of their return trips and need to be certain that they do not get stranded.

Limitations in transit services and routes are almost always necessary for reasons of cost-effectiveness. The low ridership levels that would be generated on many routes simply cannot justify the cost of providing services at these times. However, from the customers' point of view, having service at all hours and on all days is desirable. A straightforward customer-oriented measure of this aspect of service frequency is the hours per day and days per week that transit service is available for each route.

The second interpretation that customers have of service frequency is how often buses and trains come when the route is being operated. This can be measured most directly by the wait time that customers experience. When service headways (the time between successive trains or buses) are relatively short, wait time can be assumed to be one-half the headway. As headways get longer and people begin to arrive for specific scheduled trains or buses, wait times level out. However, the general inconvenience of having only a few buses or trains from which to choose continues to increase as headways are increased. Since headways and wait times usually vary by time of day and between weekdays and weekends, measuring them for customers' actual time of travel is likely to greatly improve the relationship between customer ratings and the service measures. Therefore, bus and train headways can be used as straightforward measures of service convenience reflecting the frequency of service by route, time of day, and day of the week.

Table 10.3
Similarities and Differences Across Transit Systems

	CTA Overall	CTA Red Line	CTA Blue Line	GLTC	Sun Tran
Frequent service so that wait times are short	x	x	x	x	x
Reliable trains/buses that come on schedule	x	x	x	x	x
Explanations and announcement of delays	x	x	x	x	
Trains/buses that are not overcrowded	x	x	x	x	
Freedom from the nuisance behaviors of others		x		x	x
Smoothness of the ride and stops	x	x	x	x	
Cost effectiveness, affordability, and value	x	x	x	x	
Availability of seats on the train/bus	x		x	x	
Frequency of delays for repairs/emergencies	x		x		x
Cleanliness of the train/bus interior	x	x		x	
Temperature on the train/bus	x	x			
Absence of offensive odors		x			
Friendly, courteous, and quick service			x		
Ease of paying the fare			x		
Frequency of service on weekends					x
Hours of service during weekdays					x
Short wait time for transfers					x
Connecting bus service					x
Availability of shelters and benches					x
Posted minutes to next train/bus					x
Cleanliness of stations/bus stops				x	

In addition, customers making trips that require one or more transfers are likely to view the frequency of the second and subsequent routes or services as especially important because those frequencies will dictate the amount of time that the customers can expect to spend making transfers. Transfer time is usually considered to be particularly onerous by transit riders. For this reason, it is recommended that measures of the time spent transferring are developed at least for the most important origin-destination pairs in the area served by transit.

The frequency of service is the primary determinant of actual customer wait times and one of the most important determinants of their level of satisfaction with transit service delivery. Closely related to service frequency (in customers' minds) is service reliability — the ability to stay on the expected schedules. The next section discusses this aspect of service.

10E. Reliability of Transit Service

The large number of transit agencies reporting measures of service reliability reflects the importance of providing reliable and predictable service both from a transit operations and a transit rider's perspective. Furthermore, the variety of definitions of on-time reliability reflects the different perspectives of transit agencies in measuring this service attribute (Appendix G).

It is highly advantageous both to operators and customers to maintain consistent and predictable service on transit routes and lines. For operators, a lack of regularity and uniformity leads to the inefficient use of resources (with some vehicles overloaded while others are underutilized), increased costs, and lower systemwide productivity. Two-thirds of transit operators view maintaining reliability as very important element of transit service delivery.[17] For customers, non-uniform and inconsistent service increases the level of uncertainty and uneasiness they feel at stops and stations, exacerbates crowding on vehicles and at station and stop waiting areas, and makes transfers more difficult and time-consuming.

The reliability of transit service is most often measured by on-time performance, which reflects how closely the delivery of transit service matches the published schedule. Specific measures of on-time performance include:

- percent of trains or buses reaching pre-specified points on time in different time periods, where on time is defined as arriving in a pre-specified time window;
- variance in travel times between two points;
- average minutes of bus or train delay measured at specific locations; and
- regularity of service (schedule adherence) at specific locations.

There are certain dimensions to on-time performance that make its measurement complicated. The objective of a transit rider is to arrive at his/her destination on-time, regardless of any en-route schedule variations. It is possible for trains or buses to be badly off schedule, and still get a passenger to the destination at the desired time. At the same time, transit riders are interested in minimizing the time spent waiting for vehicles since it is a component of travel time that is perceived as more onerous than invehicle travel time. It is also possible for the on-time performance measures to poorly conform to riders' experiences in this regard.

In analyzing on-time performance measures, it is often difficult to compare different types of services and different types of routes. Most on-time performance measures will have disparate ranges for different transit modes because the modes are affected by different exogenous factors. For instance, it is quite difficult to meaningfully compare the on-time performance of a commuter rail line with that of an urban bus because the bus is more vulnerable to weather problems and highway incidents. Riders recognize the inherent reliability differences, and usually customer satisfaction levels will be based on different levels of expectation.

Even within mode comparisons are difficult. To facilitate the assessment of on-time performance a distinction needs to be made between frequent transit service that is offered in small regular intervals and infrequent service that is provided according to a published schedule. In addition, the length of the route is likely to skew on-time performance results.

Because of these difficulties in comparing on-time performance for different services, it is also difficult to develop meaningful systemwide on-time performance measures. The most effective measures are obtained for specific services or small groups of services. They are best analyzed through comparisons over time as opposed to comparisons with each other.

There are also a number of operations measures that can be used as a surrogate measure for transit reliability. These measures are supply-driven and reflect the ability of the transit agency to provide the required amount of service rather than the quality of service. These measures could be used as surrogate indicators in cases where there is no option for additional data collection and analysis and include:

- the frequency of service breakdowns which is usually expressed as the average number of miles between breakdowns including a vehicle failure, road call, or service interruption, and

- vehicle availability which measures the number of vehicles that are available for service suggesting that the likelihood that service will be delivered as scheduled decreases with fewer available vehicles.

10F. Explanations and Announcement of Delays

For transit riders, one of the most difficult aspects of delays in service is the associated uncertainty about what has happened and how long they will need to wait for a train or bus. Riders are much more accommodating of delays when they are provided with information regarding the reasons for the delay and the likely length of the delay. The information allows riders to better plan ahead, and at a broader level, it helps to make riders feel like the transit system recognizes that the delays are a problem and that the transit workers are actively working on the problems.

A number of transit systems try to provide delay information to riders through on-board and station public address systems. In addition, some agencies have experimented with providing electronic map information on-board vehicles, at stations, and at bus stops. Automated Vehicle Location (AVL) systems allow operators to post real-time or close-to-real-time information for passengers.

In Europe, many transit agencies pride themselves on passenger amenities, especially the provision of customer information.[18] In London, where uncertainty about delays is among the most common sources of rider dissatisfaction, arrival time and destination information is beaconed to transit stops. In Stuttgart, the transit agency makes use of their AVL-based transit fleet management system to provide traveler information at terminal kiosks and through an in vehicle route guidance system.[19]

In addition to the more high-tech communications devices, transit agencies also provide likely-delay information to passengers through newsletters, flyers, and telephone customer service representatives.

A number of measures can be used to gauge how well delay information is being disseminated to riders, including:

- availability of on-board and station public address systems;

- availability of other electronic real-time displays;

- frequency and clarity of announcements and messages;

- percentage of significant delays for which correct information was provided to passengers on-board affected vehicles;

- percentage of significant delays for which correct information was provided to passengers waiting at affected stations or bus stops; and

- percentage of foreseeable delays (construction, maintenance, weather-related) of which customers are made aware.

Transit agencies also commonly measure the quality of their customer communications that are not directly related to delays. Some agencies reported measures that are aimed at quantifying each of the different communication efforts that transit agencies carry out. Examples of such measures include the percentage of calls by the public answered within 90 seconds; the number of service requests received by the public; and the number of calls received asking for transit-related information.

The number of complaints expressed by transit passengers is used by some agencies as a surrogate of service performance and is often reported on a monthly basis. This measure presents an effort by the transit agencies to be responsive to their clients' needs and wants. Agencies collect and analyze complaints by type (e.g. facilities, operators) and by mode and normalize the frequency of complaints by dividing by the number of transit riders or the number of transit service miles provided.

10G. Crowding On-board Trains and Buses

A common complaint about public transit systems in large cities is that trains and buses are often too crowded. Generally, the most common reasons that vehicles get overcrowded is that there is a service frequency or reliability problem, so the fact that crowding is of importance to survey respondents reinforces the importance of measuring frequency and reliability.

The crowding on-board trains and buses is an easily quantifiable measure through the calculation of various load factors. The load factors reflect the discrepancy between the available transit capacity and the corresponding transit ridership. Load factors can be expressed as the number of passengers on a vehicle divided by the vehicle's design capacity, the number of passengers divided by the crush capacity of the vehicle, or the number of passengers on a vehicle divided by the number of available seats. Passenger loading estimates are best obtained through direct observation of vehicles passing prespecified points (usually the maximum loading points).

10H. Behavior of Other Riders

Security concerns are an important element of customer satisfaction. In the surveys, these concerns manifested themselves as concerns about the behavior of other riders. If transit customers perceive that the nuisance behavior of other riders is tolerated, then their level of concern about their personal security will increase. Where there is a high level of so-called "quality-of-life" crimes and rules violations, there is more of a feeling that there is no one in charge of the system.

One way to measure the level of nuisance behavior is to track police arrest and citation records. The weakness of this approach is that it is confounded by changes in the level of effort by police to enforce system rules and by the general presence of police within the system. The presence of police officers within the system will tend to shift crimes and incidents to different places in the system, so measured improvements may not accurately reflect riders' experiences.

Some transit agencies have tried to obtain measurements on the amount of nuisance behavior by discretely sending observers into the system to collect information on fare evasion and other minor crimes and rules violations. OC Transpo in Ottawa has developed Transecure, a neighborhood watch program within its system to allow police to locate and respond to bad behavior or suspicious activities. Information from such a program is likely to be better than arrest or citation data because those observing the bad behavior will not be recognized as police. If a system is able to spend enough resources to obtain a statistically significant sample of time periods and locations, then changes over time can be monitored and compared to survey results.

10I. Smoothness of the Ride

The smoothness of the ride and the stops is an indicator of rider comfort that is not easily quantified. Smoothness can be measured on a subjective basis by having transit staff ride transit vehicles that are in operation and to rate the ride quality. Alternatively, scientific instruments could be used to measure the forces being experienced by riders as the vehicles traverse their routes.

These measures are more difficult to use and interpret than other measures discussed in this chapter. A number of factors contribute to the relative smoothness of the transit ride, including:

- the condition of the railroad track or the roadway;
- the operating condition of the rail or bus vehicles;
- the speed of the bus and the composition of the roadway traffic; and
- the experience of the rail and bus operator.

Riders' dissatisfaction about the smoothness of the trip can be caused by problems related to any or all of these factors. Therefore, developing direct measures to quantify smoothness will not necessarily help a transit operator determine whether or how to make improvements to the system to improve customer satisfaction. Given this problem, it is probably unlikely that smoothness measures would be helpful to transit operators unless they were specifically designed to isolate the different factors that go into ride smoothness.

10J. Cost Effectiveness, Affordability, and Value

The cost of travel by transit is almost always subsidized by local, state and/or national governments in an effort to promote transit use, alleviate roadway congestion, and improve the mobility of the transit-dependent segments of the population. However, in almost all cases the users are required to pay fares to use transit systems. Fare levels affect customer satisfaction and ridership.

For any given customer, the measure that is directly related to the questions of cost effectiveness, affordability, and value is the cost per transit ride. Because most systems offer some type of discounted multi-ticket fare as an option to a one-way ticket, the cost per ride may be different depending on the ticket type that individuals use. If monthly passes or another type of unlimited ride ticket types are available, the cost per ride will also vary based on the amount of system usage.

In most cases, the average cost per ride that individuals pay will vary by traveler market segment because ticket type choice will vary by market segment. Developing separate measures for different traveler market segments may be the best way to relate customer satisfaction with transit fare levels.

10K. Availability of Seats

Availability of seats is a special case of crowding on transit vehicles that is discussed above under section 10G. One can measure the ratio of the number of people on a vehicle to the number of seats on a vehicle to quantify the availability of seats.

10L. Frequency of Delays due to Repairs/Emergencies

The paramount importance of delays and reliability to transit passengers was discussed above under section 10E. However, the analysis of the survey results suggests that riders do not consider all delays equally. Delays that are due to breakdowns or accidents are particularly irksome to transit riders because they are to some extent preventable. Weather-related delays, while certainly viewed negatively, have a lesser impact than delays due to bus or train mechanical problems.

Transit agencies commonly quantify the reliability of transit vehicles with the measures mean distance between failures (MDBF) or average failures per vehicle. Operations staff use these measures to detect problems with vehicles of one type or another, so separate values are calculated for each vehicle type in the system. The primary advantage of these measures is that most agencies collect this information on a continuing basis anyway, so no additional data collection is necessary.

The primary disadvantage of these measures is that they are not collected for the purpose of measuring the quality of customer service delivery. To relate the measure to particular riders' customer satisfaction, it is sometimes necessary to obtain detailed information about the vehicle types being used on specific routes and to calculate route-specific or service type-specific weighted averages of the mean distance between failures. In addition, the type and circumstances of failures will have a large impact on customers' perceptions and this information is not necessarily captured by the maintenance measures. It would probably be quite useful to categorize the specific problems causing the breakdowns, whether or not passengers were able to be rerouted once a vehicle broke down, and the response time to address the incident.

The frequency of transit-related accidents was another category of measures cited by many agencies. Some of the agencies normalize the number of accidents per miles of service while other agencies break out accidents by type including passenger accidents, employee accidents, preventable accidents, vehicle accidents, etc. Measures of accident incidence are usually reported on a monthly and a mode-specific basis.

10M. Passenger Environment On-board Vehicles and at Stations/Stops

The general environment through which passengers travel on transit has a great deal to do with their level of satisfaction. However, it is difficult to develop a consistent and objective approach to measuring the quality of the passenger environment.

Some agencies employ professionals whose responsibilities include monitoring the system from the customer's point-of-view. These individuals are trained to consistently rate stations and vehicles according to specific objective measures or on qualitative pre-set scales. This information is then

aggregated and tracked over time to measure how the passenger environment changes. The information is shared with the operations managers who are responsible for the specific elements being evaluated, so that they are able to evaluate the quality of their departments' service delivery.

New York City Transit uses its passenger environment survey to obtain data on a wide range of subway categories[20], including:

Station

- lighting at different locations within stations;
- public address system clarity;
- condition of escalators and elevators;
- presence and readability of system maps in the stations;
- amount of litter on the platforms and track bed;
- amount of stains and spills on the platforms;
- amount of graffiti in the station;
- quality of the station signage;
- condition of public phones;
- condition of turnstiles, gates, token vending machines;
- courtesy and appearance of token booth personnel;
- availability of maps and system information in the station.

Subway Cars

- exterior graffiti;
- condition of doors;
- lighting;
- air conditioning, fans, car temperature;
- clarity of station stop and safety announcements;
- amount of litter, spills, and stains in the car and;
- presence of graffitied, scratched, and cracked windows;
- appearance of guards.

Bay Area Rapid Transit (BART) performs a similar quarterly review of its facilities.[21] The BART survey includes 31 specific measures that are organized around organizational areas of responsibility. The BART measures include:

Facilities Management

- Station Patio Cleanliness
- Parking Lot Cleanliness
- Landscape Appearance

Station Operations

- Station Cleanliness
- Station Graffiti
- Restroom Cleanliness
- Advertisements in Stations
- Brochures in Kiosks

Station Agents

- Agent Available or Sign in Place
- Agent in Uniform
- Agent wearing Name Badge

BART Police

- BART Police Personnel in Stations
- BART Police Personnel in Parking Lots/Garages
- BART Police Personnel on Trains

Public Address Announcements

- P.A. Arrival Announcements
- P.A. Transfer Announcements
- P.A. Destination Announcements

Rolling Stock

- Train Exterior Graffiti
- Train Doors Operative
- Train Interior Graffiti
- Train Interior Cleanliness
- Train Window Etching
- Temperature on Trains
- Advertisements on Trains

Elevator/Escalator Availability

- Station Elevator Availability
- Escalator Availability - Street
- Escalator Availability - Platform

Automatic Fare Collection Availability

- Fare Gate Availability
- Ticket vending Machine Availability

On-Time Performance

- Train on Time
- Customer on Time

A number of the passenger environment measures are subjective and qualitative. The careful training of observers and tests to ensure that ratings are being made consistently are essential for the data collection effort to be effective. However, despite the difficulty in establishing and monitoring the data collection effort, passenger environment surveys are probably the best way for transit agencies to understand their systems from customers' perspectives.

ENDNOTES

[14] European Conference of Ministers of Transport, Round Table 92: *Marketing and Service Quality in Public Transport,* Organization for Economic Cooperation, Paris, France, 1993.

[15] Transit Cooperative Research Program, Research Results Digest, Number 4, *Transit Ridership Initiative*, Transportation Research Board, National Research Council, Washington D.C., February 1995.

[16] Transit Cooperative Research Program, Research Results Digest, Number 3, *Total Quality Management in Public Transportation*, Transportation Research Board, National Research Council, Washington D.C., October 1994.

[17] National Cooperative Transit Research & Development Program, Synthesis 15, *Supervision Strategies for Improved Reliability of Bus Routes*, Transportation Research Board, National Research Council, Washington D.C., September 1991.

[18] Transit Cooperative Research Program, Research Results Digest, Number 22, *International Transit Studies Program - Report on 1996 Missions*, Transportation Research Board, National Research Council, Washington D.C., October 1997.

[19] Transit Cooperative Research Program, Research Results Digest, Number 20, *International Transit Studies Program - Report on the First Three Missions*, Transportation Research Board, National Research Council, Washington D.C., May 1997.

[20] Charles River Associates Incorporated, *Metropolitan Transportation Authority Comprehensive Line Improvement Study*, March 1994.

[21] Aaron Weinstein and Rhonda Albom, *Securing Objective and Reliable Data on the Quality of the Passenger Environment — The Redesign of BART's Passenger Environment Measurement System,* presented at the 77th Annual Meeting of the Transportation Research Board (January 1998).

This page left intentionally blank.

CHAPTER 11. AN OVERVIEW OF DATA COLLECTION AND ANALYSIS METHODS

In this chapter we outline the broadly defined desirable features of a data collection and analysis plan. The differences in the level of service offered and the nature of the markets served by each transit system do not allow the development of a unique set of specific data collection and analysis procedures. Furthermore, the identification of a different set of priorities for service improvements by riders of different transit systems further stresses the need for a customized approach to data collection and analysis.

The broadly defined principles guiding the data collection and analysis approach are presented in two sections. We first outline the elements of a data collection plan that minimizes biases and aggregation errors, provides data that are internally consistent and relevant from a passenger perspective, and accounts for the statistical significance of the collected data at a reasonable cost. We conclude our discussion by outlining different ways of analyzing the collected transit performance data and summarizing the results.

11A. Principles of Data Collection

In order to gauge the quality of customer service by measuring specific service attributes, it is essential that the transit agency consider the quality of the data that are being collected and the appropriateness of the chosen data collection method(s). As noted in the previous chapter, data on different service measures can be obtained by a variety of manual and automatic methods.

The manual methods include observation of service attributes by field inspectors, by field worker data collection staff, and by "mystery riders," transit agency staff or contractors who ride the system as customers would without letting transit workers know who they are or where they will be. In many cases, inspectors assemble the data that would be used in evaluating service attributes for their own purposes, thus the added cost of using this information for customer service evaluation is low. Special data collection procedures by transit staff and mystery riders can be used to obtain the other service measures.

Some transit service measures can be recorded automatically. For instance, systems that use buses equipped with AVL systems can automatically collect data on vehicle headway, on-time performance, and ridership allowing us to calculate a multitude of performance measures discussed in this report. Furthermore, the implementation of an AVL system allows the development of passenger information systems that can be used to provide estimated time of arrival to waiting passengers, display vehicles on an electronic map at a bus stop or rail station, and provide en route information to transit passengers.

A review of the current status of AVL bus transit systems in the U.S. along with a detailed technical review of different AVL architectures and technologies is presented in a recent TCRP report.[22] The advantage of such a data collection mechanism is that a variety of accurate performance data can be automatically collected at the route level by time of day and day of the week. At the same time, the challenge with these data is the ability to properly sample, organize, and analyze the information that is gathered in order to obtain the meaningful measures that are being sought.

Planners need to be aware that there are several potential problems with any given measure that can reduce its usefulness in analyzing service delivery. Among the potential problems are:

- bias;
- aggregation error;
- inconsistency;
- irrelevancy form the passenger perspective;
- insignificance; and
- cost to assemble and analyze data.

These issues are discussed below.

Bias. In this context, bias refers to a set of systematic errors that tend to overstate or understate the performance of the system for a specific measure. Performance measures should be as free from bias as possible. Examples of biased measures include data from a non-representative sample of routes or services and data assembled with methods that cause the observed situation to be different than that experienced by riders. If an agency were to assess the reliability of its bus system by measuring on-time performance only on routes of one type, say routes that serve major downtown stops, erroneous conclusions about the system as a whole are likely. Similarly, if an agency were to evaluate aspects of customer service by having uniformed inspectors observe transit employees' interactions with customers, then it is likely the results of such an evaluation would not reflect conditions when inspectors were not present.

Aggregation Error. If service measures are collected at too gross a level, important nuances of customer service delivery will be lost. For instance, if on-time performance was calculated on a systemwide basis and was used to gauge customer satisfaction with on-time reliability, it is possible that the measure is masking significant differences between different routes and lines. If a small number of routes have significantly poorer performance than the system as a whole, their effect on the objective service measures will understate the negative effect that they have on customer satisfaction.

Inconsistency. Because the most effective way to analyze service measures is to analyze changes over time and differences between different routes and services, the measures of service delivery and the scales used to record them should be consistent over time, from location to location, and from one evaluator to another. This is particularly important for the more subjective measures such as cleanliness. If inspectors or members of the field staff are employed to rate stations or vehicles on cleanliness, each one of them should have consistent ratings. In addition, the ratings should not vary with time. This is sometimes difficult because changes in the level-of-acceptability of certain conditions are likely to occur over time, particularly if a system invests in improvements in the specific aspect of service under study.

When agencies employ staff to make subjective measurements of service measures, the following steps should be taken whenever possible:

- develop objective measures whenever possible (e.g., use a thermometer to measure the temperature on vehicles, rather than a field worker rating of temperature);

- train the field workers extensively, employing actual field evaluations, to ensure that different fieldworkers rate things consistently;

- test inter-rater variations in ratings to ensure that raters remain consistent (sometimes the best way to test this is to have raters have some overlapping responsibilities).

Irrelevancy to Customers. Often, it is possible to use already-collected measures of performance to evaluate service delivery to customers. Of course, whenever this is possible it is desirable from an efficiency point-of-view. However, because these data are collected for purposes other than the evaluation of customer service delivery, planners need to assess the relevancy of the measure to customers. For example, information on on-time performance is commonly collected at train and bus terminals. In many cases where ridership is highly directional or is skewed to be on only part of the route or line, on-time performance at a particular terminal point may be largely irrelevant to customers. If a morning peak train runs close to schedule going into the CBD but then is delayed after it has made it past the CBD, the delay is irrelevant to the large majority of riders. In this case, a better on-time performance measure would be one that was collected at a CBD station.

Insignificance. In order to draw valid conclusions from the assessment of service measures, an agency needs to ensure that enough data are sampled and assembled to make the conclusions statistically significant. An agency should first define finite elements of its system, such as stations, buses in a particular time period, or buses on a particular route. As a second step, statistical sampling methods should be applied to determine how many of the elements need to be studied or observed in order to make statistically valid conclusions. If information is assembled in an ad hoc way, it is possible that variations in service quality will never be accurately observed.

Cost to Assemble Data. Finally, as for any primary data collection effort, the costs of getting particular types of data need to be considered and traded-off with the benefits of the data that would be collected. In general, the errors introduced by the potential problems described above can be reduced somewhat through more and better data collection efforts that almost always increase the cost of data collection. Although it is difficult to determine the cost-effectiveness of data collection efforts, the agency should set as a priority maintaining data on the measures associated with the three or four of the most important aspects of service from the customer's point-of-view.

For those aspects of service that are perceived as less important, an agency should probably obtain data through less rigorous methods, perhaps using less accurate measures that are already collected or are easily collected. In developing cost estimates for service data collection, an agency should seriously consider the added benefits of maintaining the data assembly over time, rather than on a one-time basis. In addition, an agency should consider collecting detailed high-quality data for specific elements of the system, rather than the system as a whole.

11B. Approaches to the Analysis of Performance Data

The ultimate objective of the analysis of the transit performance measures is to facilitate a focused and accurate assessment of any existing weaknesses in service and the measures that need to be taken in response to these performance problems. To provide transit management with a means of identifying the strengths and weaknesses of transit service and supporting its evaluation the analysis should, as stated earlier:

- provide transit management with a systemwide overview of transit operations for different transit modes;

- evaluate transit performance on a route-specific level of detail by focusing on individual segments of the transit network;

- monitor changes in transit service over time to identify deteriorating conditions or to highlight improvements in service in response to service intervention;

- identify the variation in transit level of transit service by collecting data specific to a service area, time of day, or day of the week for the service attributes of interest; and

- guide the development of marketing and communication strategies to inform transit customers and potential customers of the desirable service features.

To provide transit management with these insights, we demonstrate four different broadly defined ways in which the collected transit performance data can be analyzed. We use as a hypothetical example a measure of bus on-time reliability as reflected in the percentage of buses arriving late at the central business district bus terminal. We have also assumed that comparable data on on-time performance are available for four different points in time between 1979 and 1997. The figures that are presented and discussed allow us to:

- measure bus performance at a systemwide level and compare it with differences in performance at the bus route level;

- identify trends in systemwide and route-specific levels of bus performance over time;

- assess differences in the perceptions of different market segments including bus riders and nonusers, frequent and infrequent transit riders, riders using different routes, and riders with different socioeconomic characteristics; and

- compare riders' perceptions to measures of transit service to identify whether the strengths and weaknesses perceived by riders actually reflect the level of transit service that is currently provided.

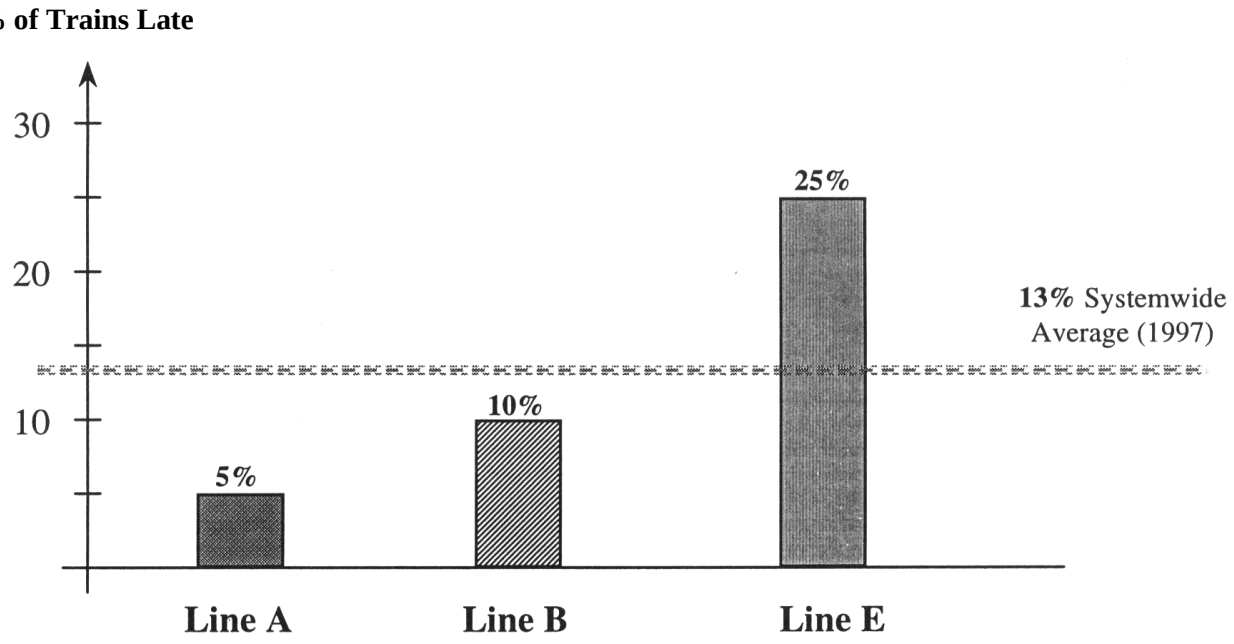
These layers of analysis correspond to an ever-increasing level of complexity. It is therefore not expected that all layers of analysis will be employed by each agency to study each of the important aspects of service. Furthermore, the more complex analyses presented below also require a wealth of data that may be maintained only for a few important measures of service.

I. Cross-Sectional Analysis of Transit Performance

The analysis of on-time transit reliability at a single point in time can provide a snapshot of transit performance both at a systemwide and at a transit route level. Although the systemwide measure can be a useful indicator of overall performance especially when monitored over time, it is also important to focus on the performance over sections of the transit system to identify potential differences by line.

Figure 11.1 presents a hypothetical example where the aggregation at the bus system level without any attention to the disaggregate route level of detail would mask important differences in performance by bus route. As shown in Figure 11.1, the overall on-time performance for the transit bus system is reflected on a satisfactory systemwide average of 87% of buses arriving within a specified time interval. However, a more detailed analysis of on-time performance at the route level suggests that there are considerable differences in route performance that would ordinarily be masked by focusing solely on the systemwide average measure.

Figure 11.1
Comparative Route Analysis



Therefore, on the basis of such a cross-sectional analysis of the bus system, the analysis would conclude that:

- the overall level of bus on-time performance is satisfactory, but
- there are important differences by route which suggest that:
  - route E experiences a significant amount of buses that are late and should be identified as a priority for service improvements;
  - route B operates at an acceptable better-than-average level but should be monitored to prevent any deterioration in service; and
  - route A should be used as a benchmark of on-time performance for the whole system.

II. Historical Analysis of Transit Performance

An additional layer of analysis can be provided by the study of systemwide and route specific on-time performance over time. Such an analysis can be used to identify trends of progress and deterioration in transit service that are not provided by the snapshot provided by the cross-sectional analysis.

A review of the hypothetical historical patterns of on-time performance for the same system shown in Figure 11.2 uncovers some important trends that could help explain the differences in on-time reliability across the system. In particular, it appears that the systemwide trend of deteriorating on-time

performance has been reversed in the past three years. However, there are some important underlying differences among the three routes suggesting that:

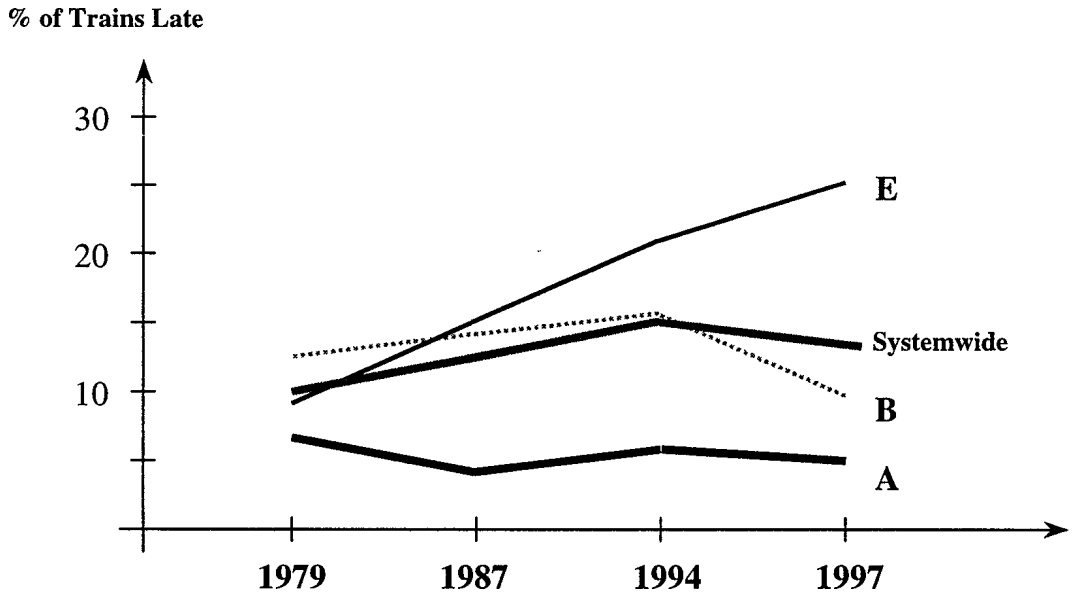
- the current poor on-time performance for route E is the result of an ongoing deterioration in transit level of service and reflects problems that date back more than a decade and that have gradually affected transit service;

- route B has enjoyed improved on-time reliability over the past three years reflecting the systemwide trend; and

- route A has maintained an excellent level of service over time.

Thus, despite the improvement in systemwide service performance the identified route-specific patterns of stability, progress, and deterioration in service performance over time can be used to support route-specific interventions.

Figure 11.2
Performance Monitoring Over Time



III. Riders' Attitudes and Transit Performance

The third layer of analysis that supplements the cross-sectional and historical analysis of transit performance data focuses on the joint analysis of transit riders' attitudes and transit performance. Two general types of analysis can be accommodated within this context. First, an analysis of the differences in attitudes across segments of the transit market can help identify opportunities for marketing to different groups of riders. Second, a comparison of attitudes and transit performance can help identify riders' misperceptions and identify opportunities for communicating service improvements to transit riders.

Figure 11.3 illustrates the differences in perceptions among users and nonusers as reflected on their ratings of five different aspects of transit service. A rating scale of 0 to 10 was used with higher values corresponding to more positive perceptions of transit service. As shown in Figure 11.3, current transit riders rate all aspects of transit service, with the exception of safety while using the system, higher than nonusers do. The pattern of differences in the respondents' ratings suggests that:

- the transit agency needs to allocate resources to enhance riders' perception of feeling safe and secure while riding the transit system;

- the perception of safety and security among nonusers does not appear to be the primary reason for not using the transit system;

- the gap between users' and nonusers' perceptions is greater for "frequency of transit service" and "transit on-time performance" which are perceived rather positively by current riders of the transit system; and

- there are considerable opportunities to improve nonusers' perceptions of transit service along most of the dimensions of transit service as part of an effort to increase transit ridership.

Although the example of such an analysis is presented at the systemwide level for transit riders and nonusers it can be further expanded along two additional dimensions. First, route-specific analyses can be conducted for routes and groups of routes that are of greatest interest to the transit authority. Second, comparisons of attitudes among market segments can be expanded to account for differences among frequent and infrequent riders, male and female riders, and riders with different degrees of captivity to transit. These analyses can provide insight into the appeal of different transit routes to distinct segments of the market.

Finally, it is possible that the availability of transit performance and survey data at similar points in time allow comparisons between riders' perceptions and transit performance measures. Such comparisons are again most meaningful if they can be repeated over time and across different routes of the system. The availability of such data supports a fourth layer of analysis that can be used to relate patterns of change in transit performance to changes in riders' perceptions.

The comparisons that can be made allow us to identify cases where service improvements have a positive impact on riders' perceptions and cases where despite improvements in transit service transit riders' perceptions continue to remain rather low.

Figure 11.3
Perceptions of Users and Nonusers for Route A

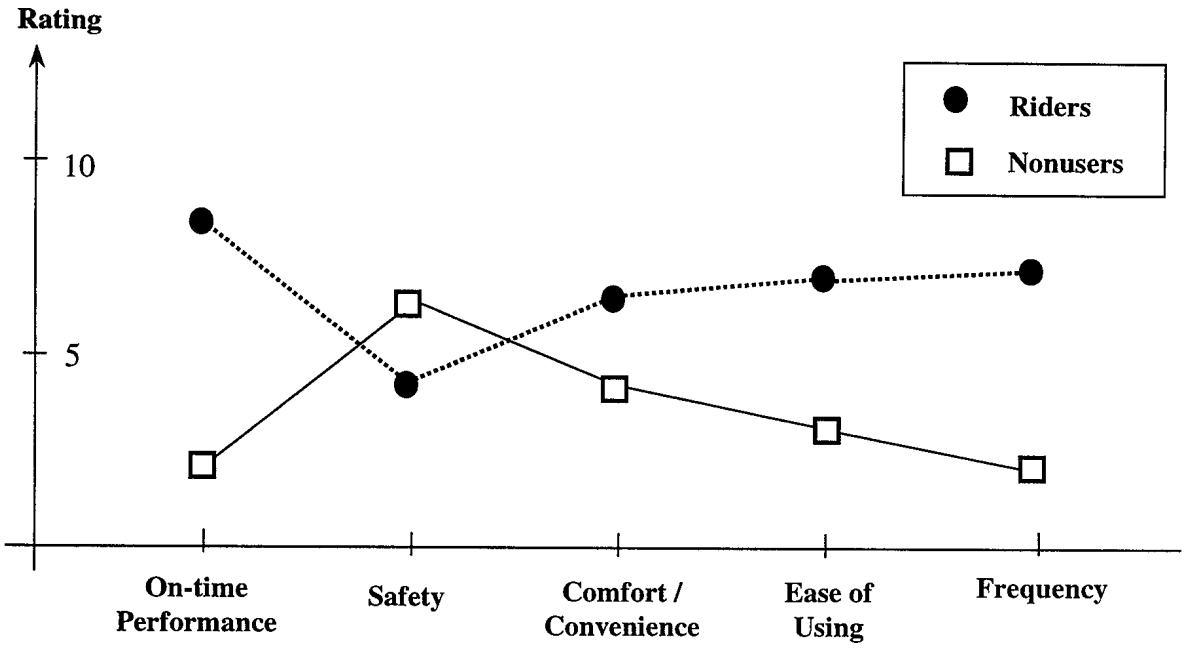


Figure 11.4 offers an example of comparisons that can be made using historical attitudinal data and corresponding performance data at the route level to identify the extent to which there is a correlation between traveler perceptions and transit performance.

The bar chart and the left hand axis illustrates the average ratings given by riders of routes A and E on a scale of 0 to 10 with higher values corresponding to more positive perceptions of service. The line graph and the right hand axis correspond to the on-time performance reflecting the percentage of buses arriving late for the A and B routes at the three study years.

The comparisons that can be made suggest that:

- riders' ratings for route E are consistently lower than those by riders of route A properly reflecting the historically better on-time performance of route A;

- route E riders' ratings of the transit service have dropped over time in a manner that is consistent with the deteriorating performance of route E;

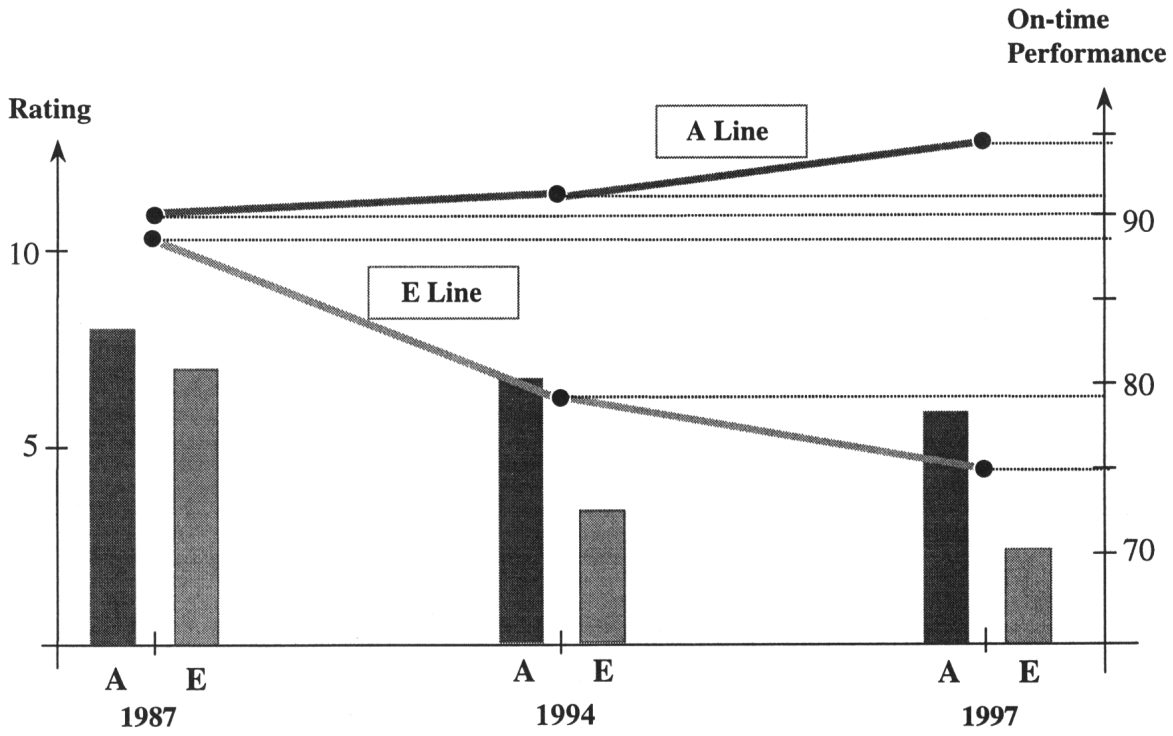
- the gap between the ratings for route A and E has widened over time again properly corresponding to the widening gap in the level of transit on-time performance offered by each route; and

- the drop over time in riders' ratings of route A is not consistent with the high level of on-time performance for route A.

These observations suggest that riders' evaluations are generally consistent with the level of service that is provided. The need to improve the on-time performance along route E is supported both by the existing low level of on-time reliability on that route as well as the low ratings provided by riders. It is expected that the implementation of such service improvements will enhance route E riders' perceptions and bring them closer to the ratings provided by riders on route A.

Finally, the apparent inconsistency between the historically high level of on-time reliability for route A and the steady or decreasing ratings by route A riders suggests that other aspects of the performance for this route need to be examined more closely. It is possible that due to deterioration in other service characteristics for route A, riders provide ratings for on-time reliability that are lower than expected. However, if there are no apparent weaknesses in other aspects of route A service, the implementation of a marketing campaign aimed at riders of route A may be considered to stress the existing high level of service.

Figure 11.4
Performance Measure versus Riders' Perceptions



ENDNOTES

[22] Transit Cooperative Research Program, Synthesis 24, *AVL Systems for Bus Transit*, Transportation Research Board, National Research Council, Washington D.C., 1997.

APPENDIX A

CUSTOMER SATISFACTION/DISSATISFACTION RESEARCH — AN HISTORICAL PERSPECTIVE

Consumer behavior as a distinct discipline dates only from the mid 1960s. Interest in understanding and tracking specific consumer problems grew dramatically in the late 1970s under the broad label of consumer satisfaction/dissatisfaction (CS/D) research. Its growth coincided with (and was abetted by) a growing interest on the part of both government regulators and leaders with the consumer movement in making the policy formulation process more rational and systematic. Critics of past consumer policy formulation had argued that it was too often influenced by chance events, letter-writing campaigns, media publicity, and partisan political agendas. The earliest comprehensive CS/D studies were, in fact, motivated by the policy planning needs of a public regulatory agency, the Federal Trade Commission (Technical Advisory Research Program (TARP) 1979), and a private non-profit sector organization, Ralph Nader's Center for Study of Responsive Law.

Pioneering studies by Handy and Pfaff in the mid 1970s developed raw and weighted indexes of consumer satisfaction with food products across seven broad food categories. After that point, research on the topic grew rapidly.

Since 1985, two different patterns have emerged. First, there has been a considerable drop in CS/D research from a public policy perspective. At the same time, however, there has been substantial growth in interest in the topic of consumer satisfaction research within the private sector. This has been driven primarily by the growth of the service sector of the economy where managers have realized that tracking satisfaction is crucial to success when intangibles such as personal attention and atmospheres are the "product." A number of private satisfaction tracking services have emerged. Many of these services have made extensive use of earlier methodological developments in social policy research.

Initial studies on CS/D sought to calibrate the amount and types of dissatisfaction in the marketplace as a basis for policy planning. This body of research was largely descriptive (TARP 1979). Wide variation was found across purchase categories. These studies differ widely in the basic measure of dissatisfaction they used. Some focused on more or less objective measures of "problems," others on subjective feelings of "dissatisfaction." Some counted any negative experience whatsoever, some only "serious" dissatisfactions, and some only the most recent problem. Also, there was the issue of opportunity for problems. Measures did not always control for frequency of purchase. Definitional problems persist today.

Most of the early studies were based on survey data. An alternate approach was complaints data, data on the extent to which consumers voluntarily speak up about their dissatisfactions. Such data have the advantage of not requiring field surveys; however, they are typically biased in two important ways. First, some types of problems in some types of industries are more likely to be voiced than others, and some problems are less serious than others, and or less costly than others. Monopolies are often relatively "immune" to complaining except from a small elite. Still other industries are more encouraging of complaints. Finally, not all consumers complain. These problems have led researchers in recent years to fall back on the more costly, but more objective, survey research methods.

Finally, most CS/D research from 1975 to 1985 was conducted within product and goods producing industries. Only after 1980 were initial concepts and models developed to measure consumer satisfaction/dissatisfaction within service industries.

LITERATURE SEARCH SUMMARY FOR SERVICE QUALITY AND CUSTOMER SATISFACTION MEASUREMENT — OUTSIDE TRANSIT INDUSTRY

***Conceptual Model of Service Quality and Its Implications for Future Research*, A. Parasuraman, Valerie A. Zeithaml, and Leonard L. Berry, Journal of Marketing, Fall 1985, Vol. 49, Number 4, pp. 41-50.**

Overview

The attainment of quality in products and services was a pivotal concern of the 1980s. While quality in tangible goods has been described and measured by marketers, quality in services is largely undefined and unresearched. The authors attempt to rectify this situation by reporting the insights obtained in an extensive exploratory investigation of quality in four service businesses and by developing a model of service quality. Propositions and recommendations to stimulate future research about service quality are offered.

Quality and measurement are not easily articulated by consumers (Takeuchi and Quelch 1983). Interpretation and measurement of quality also present problems for researchers. While the substance and determinants of quality may be undefined, its contribution to increasing market share and return on investment is unequivocal.

Existing Knowledge About Service Quality

Knowledge about goods quality is insufficient to understand service quality. Three well-documented characteristics of services — intangibility, heterogeneity, and inseparability — must be acknowledged. Because they are performances rather than objects, precise manufacturing specifications concerning uniform quality can rarely be set. Because of intangibility, the firm may find it difficult to understand how consumers perceive their services and evaluate service quality (Zeithaml 1981).

Second, services, especially those with high labor content, are heterogeneous: their performance often varies from producer to producer, from customer to customer, and from day to day. Consistency of behavior from service personnel (e.g., uniform quality) is difficult to ensure (Booms and Bitner 1981) because what the firm intends to deliver may be entirely different from what the customer receives.

Third, production and consumption of many services are inseparable (Carmen and Langeard 1980, Upah 1980). In labor intensive services, quality occurs during service delivery, usually in an interaction between the client and front-line personnel.

Service quality literature traditionally agrees that service quality is a measure of how well the service level delivered matches customer expectations. Delivering quality service means conforming to customer expectations on a consistent basis. (Lewis and Booms 1983)

Insights from Exploratory Qualitative Investigation

A set of discrepancies or gaps exists regarding executive perceptions of service quality and the tasks associated with service delivery to consumers. These gaps can be major hurdles to attempting to deliver a service which consumers would perceive as being high quality. Figure A.1 on the following page shows the five gap areas identified.

These are:

GAP 1: *Consumer expectation — management perception gap*
Discrepancies between executive perceptions and consumer expectations. Service firm executives may not always understand what features denote high quality to consumers in advance, what features a service must have in order to meet consumer needs, and what levels of performance on those features are needed to deliver high quality service.

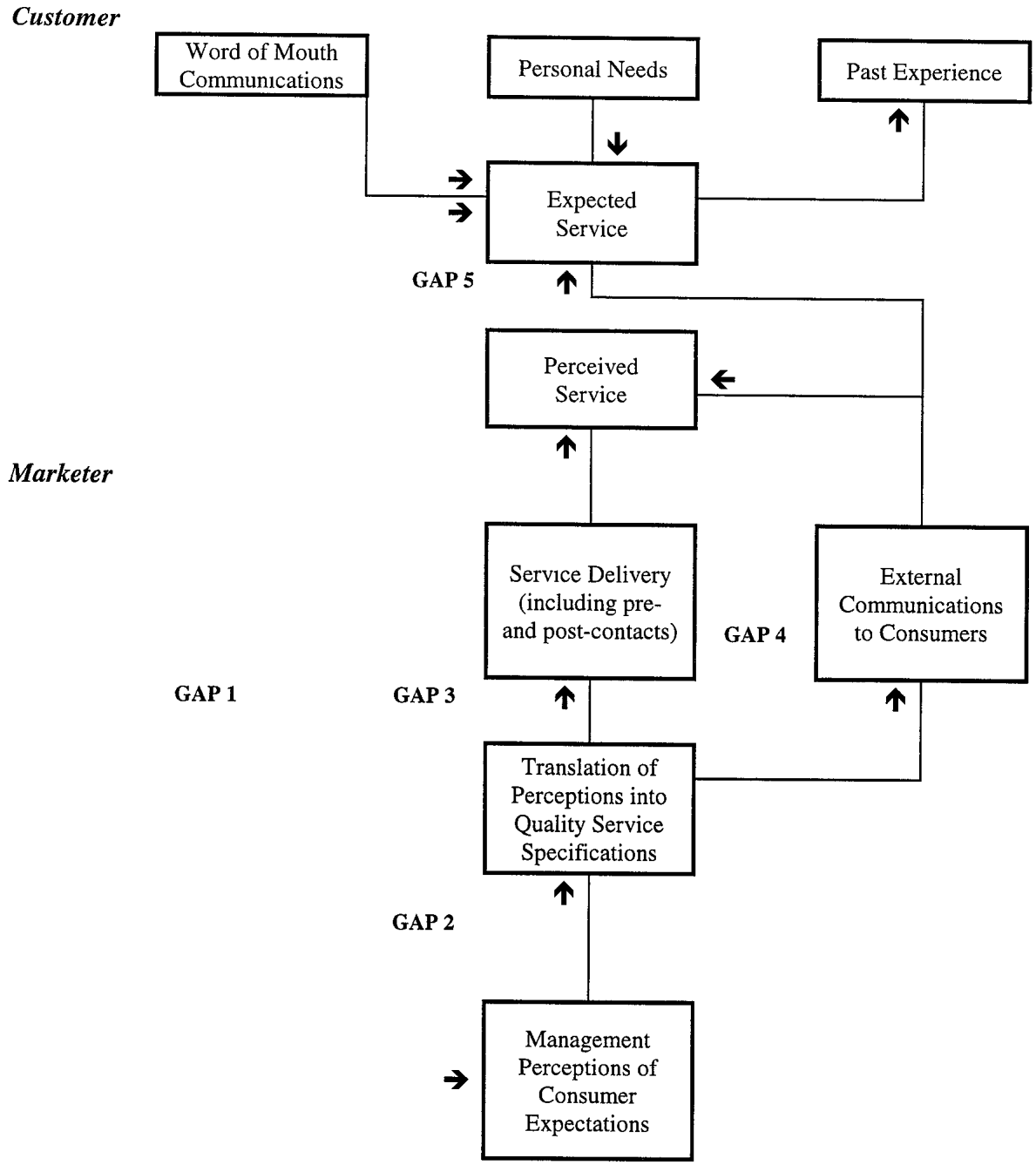
GAP 2: *Management perception — service quality specifications*
Constraints (resources, or market conditions) which prevent management from delivering what the consumer expects, or the absence of total management commitment to service quality.

GAP 3: *Service quality specifications — service delivery gap*
Difficulty in standardizing employee performance even when guidelines exist for performing services well and treating consumers correctly.

GAP 4: *Service delivery — external communications gap*
Media advertising and other communications by a firm can affect consumer expectations. Promising more than can be delivered will raise initial expectations but lower perceptions of quality when the promises are not fulfilled. Also firms can neglect to inform consumers of special efforts to ensure quality that are not visible to consumers thereby affecting consumer perceptions of the delivered service.

GAP 5: *Expected service — perceived service gap*
How consumers perceive the actual service performance in the context of what they expected. The quality that a consumer perceives in a service is a function of the magnitude and direction of the gap between expected service and perceived service.

Figure A.1
Service Quality Model



A Quality Service Model

The foundation of the model is the set of gaps shown in Figure A.1. Service quality as perceived by a consumer depends on the size and direction of GAP 5 that, in turn, depends on the nature of the gaps associated with the design, marketing, and delivery of services. The gaps on the marketer side of the equation can be favorable or unfavorable from a service quality perspective. That is, the magnitude and direction of each gap will have an impact on service quality.

The Perceived Service Quality Component

This exploratory investigation suggests that, regardless of the type of service, consumers used basically similar criteria in evaluating service quality. These criteria seem to fall into 10 key categories that are labeled "service quality determinants." These determinants are listed in Table A.2 below. Overlap among the 10 determinants may exist.

Table A.2
Determinants of Service Quality

1 RELIABILITY involves consistency of performance and dependability.

2 RESPONSIVENESS concerns the willingness or readiness of employees to provide service. It also involves timeliness of service.

3 COMPETENCE means possession of the required skills and knowledge to perform the service.

4 ACCESS involves approachability and ease of contact.

5 COURTESY involves politeness, respect, consideration, and friendliness of contact personnel.

6 COMMUNICATION means keeping customers informed in language they can understand and listening to them. It may mean that the company has to adjust its language for different consumers — increasing the level of sophistication with a welleducated customer and speaking simply and plainly with a novice.

7 CREDIBILITY involves trustworthiness, believability, and honesty. It involves having the customer's best interests at heart.

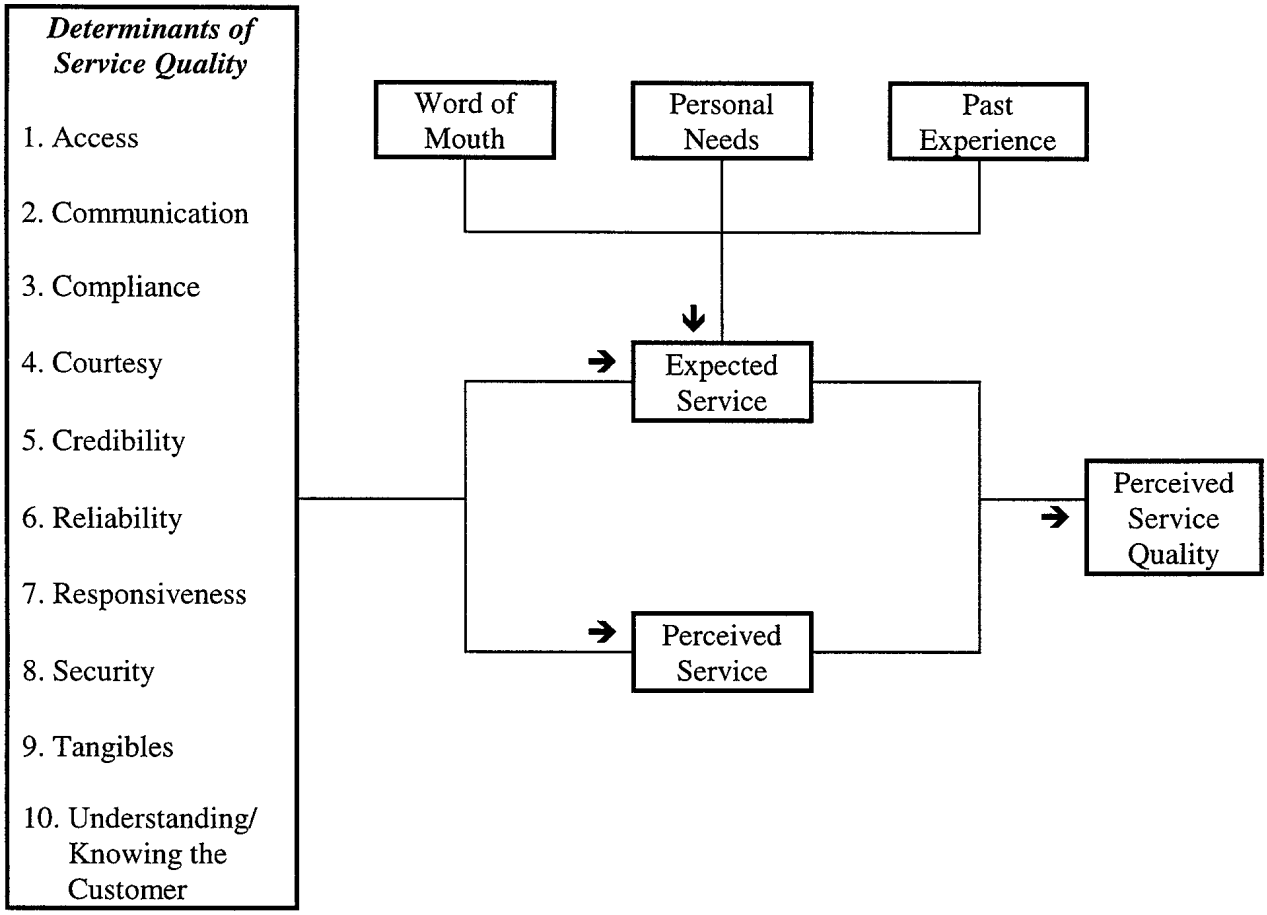
8 SECURITY is the freedom from danger, risk, or doubt.

9 UNDERSTANDING/KNOWING THE CUSTOMER involves making the effort to understand the customer's needs.

10 TANGIBLES includes the physical environment and representations of the service.

It is quite possible that the relative importance of the 10 determinants in molding consumer expectations (prior to service delivery) may differ from their relative importance vis-à-vis consumer perceptions of the delivered service. Figure A.3 on the following page indicates that perceived service quality is the results of the consumer's comparison of expected service with perceived service.

Figure A.3
Determinants of Perceived Service Quality



Two of the determinants which consumers appear to have difficulty evaluating are *competence* (the possession of the required skills and knowledge) and *security* (freedom from danger, risk, or doubt). Consumers are probably never really certain of these attributes, even after experiencing the service.

Perceived service quality is posited to exist along a continuum ranging from ideal quality to totally unacceptable quality, with some point along the continuum representing satisfactory quality. The position of a consumer's perception of service quality on the continuum depends on the nature of the discrepancy between the expected service (ES) and perceived service (PS). When ES > PS perceived quality is less than satisfactory; when ES = PS perceived quality is satisfactory; and when ES < PS, perceived quality is more than satisfactory and will tend toward ideal quality.

Although the preliminary research showed that consumers used similar criteria in judging service quality, the group participants differed on the relative importance of those criteria to them, and their expectations along the various quality dimensions. Research needs to determine whether identifiable service quality segments exist and whether and in what ways consumer expectations differ.

Takeuchi, Hirotaka and John A Quelch (1983), "Quality Is More Than Making a Good Product," *Harvard Business Review*, 61 (July-august), 139-145.

Zeithhaml, Valerie A. (1981) "How Consumer Evaluation Processes Differ Between Goods and Services," in *Marketing of Services*, J. Donnelly and W. George, eds., Chicago: American Marketing, pp. 186-190.

Carmen, James M. and Eric Langeard (1980), "Growth Strategies of Service Firms," *Strategic Management Journal*, 1 (January-March), pp. 7-22.

Upah, Gregory D. (1980), "Mass Marketing in Service Retailing: A Review and Synthesis of Major Methods," *Journal of Retailing*, 56 (Fall), pp. 56-76.

Lewis, Robert C. and Bernard H. Booms (1983), "The Marketing Aspects of Service Quality," in *Emerging Perspectives on Services Marketing*, L. Berry, G. Shostack and G. Upah, eds., Chicago: American Marketing, pp. 99-107.

Churchill, G.A. Jr. and C. Suprenaut (1982), "An Investigation into the Determinants of Customer Satisfaction," *Journal of Marketing Research*, 19 (November), pp. 491-504.

***A National Customer Satisfaction Barometer: The Swedish Experience*, Claes Fornell, Journal of Marketing, January 1992, Volume 56, Number 1, pp. 6-21.**

Overview

Many individual companies and some industries monitor customer satisfaction on a continual basis, but Sweden is the first country to do so on a national level. The annual Customer Satisfaction Barometer (CSB) measures customer satisfaction in more than 30 industries and for more than 100 corporations. The new index is intended to be complementary to productivity measures. Whereas productivity basically reflects quantity of output, CSB measures quality of output (as experienced by the buyer). The author reports the results of a large-scale Swedish effort to measure quality of the total consumption process as customer satisfaction. Efforts to measure customer satisfaction on a nationwide basis are now underway in several other countries including the U.S., Japan, and Norway.

The U.S index is the result of a joint venture between the American Quality Foundation and the University of Michigan Business School. The significance of customer satisfaction and its place within the overall strategy of the firm are discussed.

Inherent Differences Among Industry and Firm Customer Satisfaction Levels

Substantial literature suggests that market share leads to profitability (see Buzzell and Gale 1987). Customer satisfaction also is believed to lead to profitability (Business International 1990). Traditionally, much more effort is devoted to the offense for customer acquisition then to the defense to protect the present customer base (Fornell and Wernerfelt 1987, 1988). However, in the face of slow growth, a good defense is critical. Defensive strategy involves reducing customer exit and switching. One way of accomplishing this objective is to have highly satisfied customers. While improving market share and improving customer satisfaction individually result in higher profitability, it is far from certain that market share and customer satisfaction themselves are positively correlated. If an industry or company enjoys high levels of customer satisfaction, decreases in market share (perhaps because of a rise in cost) are less likely to affect profitability. Decision making in this situation is a combination of price

and quality. However, it is more difficult for a firm with a large market share to also have a high average level of customer satisfaction, especially if customer needs or wants are heterogeneous.

The ideal point conceptualization as one aspect of customer satisfaction suggests a new hypothesis about market structure and customer satisfaction. The contention is that the monopoly will have a lower score on customer satisfaction indexes than other non-monopoly industries, if it faces a heterogeneous demand. Lower customer satisfaction in this case is partially a reflection of the difficulty in serving a heterogeneous market with a limited variety of service or product offerings. On the other hand, we would expect that industries characterized by a good fit between the levels of demand and supply heterogeneity (homogeneity) to have higher customer satisfaction ratings than those with a poor fit. Industries, including monopolies, that supply a high quality homogeneous product to a homogeneous market should have high satisfaction.

Also explored is the impact of customer satisfaction on repeat business and customer loyalty in different industries. Loyal customers are not necessarily satisfied customers, but satisfied customers tend to be loyal customers. Customer switching barriers comprise a host of factors that also bring about retention. Switching barriers make it more costly for the customer to switch to another supplier or mode. Transaction costs, learning costs, loyal customer discounts, customer habit, emotional cost, and cognitive effort, coupled with financial, social, and psychological risks on the part of the buyer, all add up to switching barriers. However, previously insulated organizations become vulnerable, for they are seldom well prepared and have not made the investments in quality and customer satisfaction necessary to prevent customer exit. Low barriers and weak customer satisfaction force the company to compete on price. With high satisfaction there is less price sensitivity.

Uses of the Sweden Customer Satisfaction Barometer (CSB)

To combine premises, the proposition that evolves from the ideal-point model and the switching-barrier effect suggests that customer satisfaction should be lower in industries where repeat buyers face high switching costs and where the industry offers a homogeneous product to a heterogeneous market. With this presumption in mind, the CSB in Sweden offers the following information:

- industry comparisons
- comparisons of individual firms with the industry average
- comparison over time
- predictions of long-term performance
- *Though empirical evidence is limited, increases in customer satisfaction are generally believed to: (1) shift the demand curve upward and/or make the slope of the curve steeper (i.e. lower price elasticity, higher margins), (2) reduce marketing costs (customer acquisition requires more effort, (3) reduce customer turnover, (4) lower employee turnovers (satisfied customers affect the satisfaction of front-line personnel), (5) enhance reputation (positive customer word of mouth), (6) reduce failure costs (handling customer complaints).*
- answers to specific management questions
  (Such as the effects of overall quality and price, the impact of customer expectations, the quality increase necessary to retain dissatisfied customers, price sensitivity, switching patterns, customer complaints, and effects of word of mouth.)

Highlights of CSB Measurement

The literature on customer satisfaction/dissatisfaction suggests that satisfaction is an overall postpurchase evaluation. There is no consensus on how to measure it. Hausknecht (1990) identifies more than 30 different measures that have been used in previous research. There are three different dimensions: (1) general satisfaction (as in the studies by Moore and Shuptrine 1984; Oliver and Bearden 1983; Oliver and Westbrook 1982; and Westbrook 1980), (2) confirmation of expectations (as in studies by Oliver 1977; Swan, Trawick, and Carroll 1981), and (3) the distance from the customer's hypothetical ideal product (Tse and Wilton 1988, and Sirgy 1984). Customer satisfaction for the CSB is defined as a function of these three indicators, thus the fallibility of measures is acknowledge and taken into account.

The traditional view of satisfaction/dissatisfaction as the discrepancy between perceived performance and expectation (P-E) is not dismissed *a priori* in CSB. However, CSB measurement allows for the possibility of dissatisfaction even when expectations are confirmed (a negative correlation). For example, if low quality is expected but the product is purchased nevertheless (because of supply restrictions or price), the expectations are confirmed. Clearly, the fact that expectations are confirmed is not sufficient for satisfaction.

Presumably, customers take both price and quality into account. To avoid compounding the two, for the CSB, each was measured in the light of the other — by price (given quality) and quality (given price).

For most industries surveyed, sample frames were drawn via random digit dialing with screening for customer status. In no cases were company customer lists used as sample frames. Hence data were costly but presumably more objective.

Almost all customer satisfaction research is hampered by highly skewed distributions for satisfaction. For example, in studies ranging from shoes to medical care, more than 80% of the customers were satisfied. Only in captive markets might repeat buyers be dissatisfied in general. Skewness is a problem, but it is a statistical one. Highly skewed variable distributions do not lend themselves to conventional tests of significance and, what is equally serious, lead to downward biases in correlation analysis, low reliability, and sometimes misleading arithmetic means. In CSB, the problem of skewness was handled by (1) extending the number of scale points (usually 5 or 7) to 10 to allow respondents to make finer discriminations, (2) using a multiple-indicator approach for greater accuracy, and (3) estimating via a version of partial least squares (PLS).

CSB Results

The results of the CSB fit the reasoning presented. Overall, CSB scores are significantly higher in industries where heterogeneity/homogeneity in demand is matched by the supply. Staple foods and automobiles score at the top of the CSB; the police force and television broadcasting are at the bottom. (Transportation services were not measured as a part of the Sweden CSB.) Overall, it is noteworthy that services score lower than products, both among monopolies and among competing firms.

The Effect on Customer Loyalty

Just as price elasticity varies among firms and industries, so does "customer satisfaction elasticity." It is very important to determine how sensitive the present customer base is to satisfaction. In view of the current business emphasis on quality, one may well get the impression that quality and customer satisfaction are equally important to all firms or industries. Customer satisfaction is more important (for

loyalty) in some industries than in others. Industries with low elasticities are those in which switching costs are high (police, postal services, etc.)

The most meaningful measurement of quality is how it affects customer satisfaction. Changes in satisfaction are predictors of future performance.

Buzzell, Robert D. and Bradley T. Gale (1987), *The PIMS Principles*, New York: The Free Press.

Business International (1990), *Maximizing Customer Satisfaction: Meeting the Demands of the New Global Marketplace*, Research Report. New York: Business International Corporation.

Fornell, Claes and Birger Wernerfelt (1987), "Defensive Marketing Strategy by Customer Complaint Management: A Theoretical Analysis," *Journal of Marketing Research*, 24 (November) pp. 337-346.

Hausknecht, Douglas R. (1990), "Measurement Scales in Consumer Satisfaction/Dissatisfaction," *Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior*, 3, pp. 1-11.

Moore, Ellen M. and F. Kelly Shuptrine (1984), "Disconfirmation Effects on Consumer Decision Making Processes," *Advances in Consumer Research*, Vol. 11, Thomas C. Kinnear, ed. Ann Arbor, MI: Association for Consumer Research, pp. 299-304.

Oliver, Richard L. and William O. Bearden (1983), "The Role of Involvement in Satisfaction Processes" *Advances in Consumer Research*, Vol. 10, Richard P. Bagozzi and Alice M. Tybout, eds. Ann Arbor, MI: Association for Consumer Research, pp. 250-255.

Oliver, Richard L. and Robert A. Westbrook (1982), "The Factor Structure of Satisfaction and Related Postpurchase Behavior," in *New Findings in Consumer Satisfaction and Complaining*, Ralph L. Day and H. Keith Hunt, eds. Bloomington: Indiana University, pp. 11-14.

Westbrook, Robert A. (1980), "A Rating Scale for Measuring Product/Service Satisfaction," *Journal of Marketing*, 44 (Fall), pp. 68-72.

Oliver, Richard L. (1977), "Effect of Expectation and Disconfirmation on Post-Purchase Product Evaluations: An Alternative Interpretation," *Journal of Applied Psychology*, 62 (4) pp. 480-486.

Swan, John E., Frederick Trawick, and Maxwell G. Carroll (1981), "Effect of Participation in Marketing Research on consumer Attitudes Toward Research and Satisfaction With Service," *Journal of Marketing Research*, 18 (August), pp. 365-363.

Tse, David K. And Peter C. Wilton (1988), "Models of Consumer Satisfaction Formation: An Extension," *Journal of Marketing Research*, 25 (May), pp. 204-214.

Sirgy, Joseph M. (1984), "A Social Cognition Model of Consumer Satisfaction/Dissatisfaction," *Psychology and Marketing*, I (2), pp. 27-43.

***Expectations, Performance Evaluation, and Consumers' Perception of Quality*, R. Kenneth Teas, Journal of Marketing, October 1993, Volume 57, Number 4, pp. 18-34.**

Overview

The author examines conceptual and operational issues associated with the measurement framework defined as customer "perceptions-minus-expectations" (P-E) identified by Parasuraman, Zeithaml, and Berry (1985). The examination indicates that the P-E service gap premise is of questionable validity because of a number of conceptual problems involving the (1) conceptual definition of expectations, (2) theoretical justification of the expectations component of the P-E framework, and (3) measurement validity of the expectation (E) and revised expectation (E*) measures specified in the published service quality literature.

The P-E model and two alternative perceived quality modes that are designed to address the problems associated with the P-E model are empirically tested and the implications of the conceptual issues examined in the study and of the empirical findings are explored.

Definition Problems

Alternative definitions of expected or ideal service exist. Conceptualizing service expectation as ideal standards is a problem under each of the interpretations examined.

Classic attitudinal model point interpretation (Ginter 1974; Green and Srinivasan 1978). In these models, the ideal point is the perfect or utility maximizing level of the attribute. For example, if the attribute has a non-maximum ideal point, once the ideal point is reached "there are negative utility returns for further increases in the attribute" (Lillien, Kotler, and Moorthy 1992, p.9). Favorableness of an evaluation of an attitude object is positively related to the closeness of the object to the ideal object.

Feasible ideal point interpretation. A second interpretation of the service quality ideal standard is that it represents a feasible level of performance under ideal circumstances. However, the "feasible ideal point" conception of E is not compatible with the service quality P-E measurement specification, when finite classic ideal point attributes are involved.

Operational Definition Problems

Empirical research has identified important problems concerning the operationalization of the service expectation (E) concept. Respondents may assign unrealistically high ratings to the E response scales. Carmen (1990) questions the validity of expectation measures when consumers do not have "well-formed expectations." Research by Teas (1993) indicates that a considerable portion of the variance in responses to the E scale is because of variance in respondents' interpretations of the question being asked, rather than to variance in respondents' attitudes.

To correct respondents high ratings on E scales, Parasuraman, Berry, and Zeithaml (1990) proposed a revised expectation (E*) measure, based on ratings of the attribute's "essentialness" for excellent service. However, using the revised definition of expectation (E*), in conjunction with the P-E measurement specification, suggests that high performance on essential attributes (high E* scores) reflects lower quality than high performances on attributes that are less essential (low E* scores). This measurement result is illogical.

Results of Testing Alternative Perceived Quality Frameworks

The results suggest a considerable portion of the variance of service quality expectation measures may be because of respondents' misinterpretations of the question or the scales. The empirical testing also indicates that the inclusion of attribute weights in the P-E or other alternative frameworks does not improve the validity of the models. This result is similar to the findings of other research that indicates importance weights often do not increase, and may decrease, the predictive validity of multiattribute models (Bass and Wilkie 1973).

The conceptual and operational definition problems with the P-E "gap" framework and alternative tested models create ambiguity concerning the interpretation and theoretical justification of these perceived quality concepts.

Parasuraman, A., Leonard L. Berry, and Valerie A. Zeithaml (1990), *An Empirical Examination of Relationships in an Extended Service Quality Model*, Cambridge, MA: Marketing Science Institute.

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

Ginter, James L. (1974), "An Experimental Investigation of Attitude Change and Choice of a New Brand," *Journal of Marketing Research*, 11 (February), pp. 30-40.

Green, Paul E. And V. Srinivasan (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research*, 5 (September), pp. 103-23.

Lillien, Gary L., Philip Kotler, and K. Sridhar Moorthy (1992), *Market Models*, Englewood Cliffs, NJ: Prentice Hall, Inc.

Carmen, James M. (1990), "Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions," *Journal of Retailing*, 66 (Spring) pp. 33-55.

Teas, R. Kenneth (1993), "Consumer Expectations and the Measurement of Perceived Service Quality," *Journal of Professional Services Marketing*, 8 (2), pp. 33-54.

Bass, Frank and William L. Wilkie (1973), "A Comparative Analysis of Attitudinal Predictions of Brand Preference," *Journal of Marketing Research*, 10 (August) pp. 262-269.

***Competing Based on the Customer's Hierarchy of Needs,* Doug Schaffer, National Productivity Review (Summer 1995) pp. 9-15.**

Even when companies improve their performance, they often have difficulty achieving real competitive advantage in the face of often astounding operational improvements, since most customers just do not seem very excited. This is largely because customers have been excluded from improvement efforts to date. For companies to better perform in ways that matter to their customers, they must know why customers buy from them in the first place. This represents a shifting hierarchy of needs that requires companies to improve their performance in ways that will make their customers sit up and take notice.

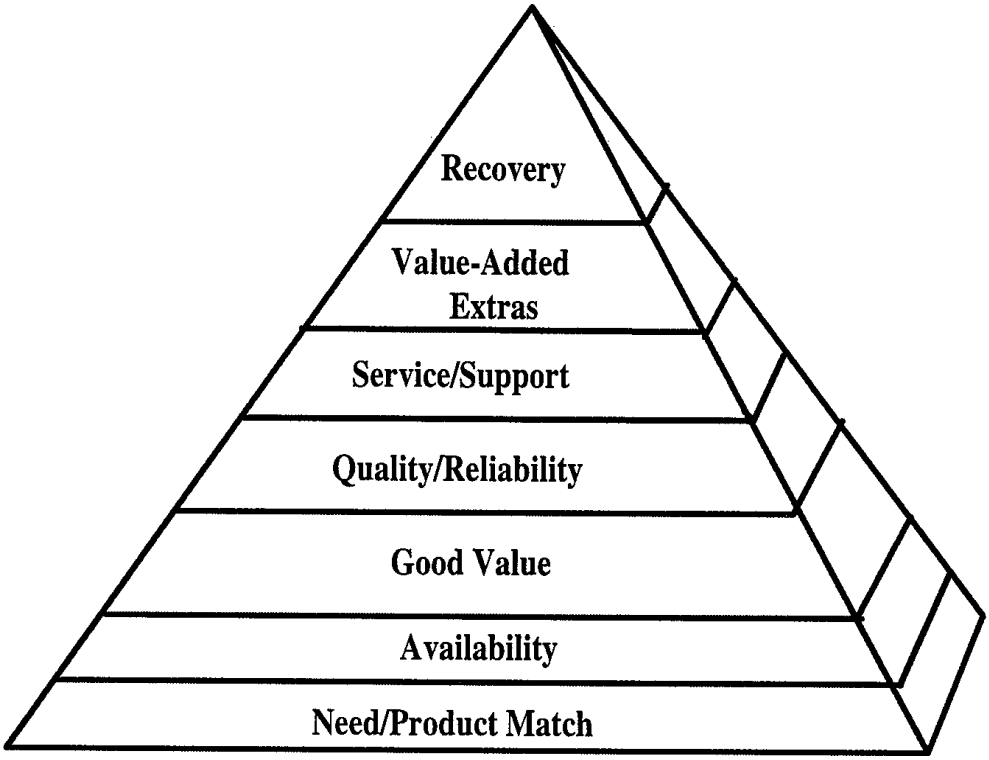
It is typical for companies to launch improvement programs in response to competitive pressures, then several years down the road report improvements that primarily affect internal operations. Published reports often list fewer engineering problems or defects, streamlined purchasing processes, lower

receivables, improved employee safety, etc. All are worthy goals and certainly contribute to a healthy balance sheet, but may be only of marginal interest to customers. Many programs to improve corporate performance are more effective in reducing costs and improving profitability than spurring growth and increasing market share.

Most companies have a rudimentary understanding of why customers buy their product or select their service. However, most would be hard-pressed to explain how much of a customer's decision is based on service characteristics, value, or reputation.

In his 1954 work, *Motivation and Personality,* Abraham Maslow proposed a theory of human motivation characterized by a hierarchy of needs. Inserting the needs of the customer into Maslow's model yields a model of customer motivation (Exhibit A.4).

Exhibit A.4
Hierarchy of Customer Needs



First on the list is how closely a product or service matches what the customer needs. The product must be available when the customer needs it. Customers expect a good value — the relationship of the cost to perceived benefit. Customers also expect quality and reliability. They never want to be stranded, inconvenienced, or endangered by products or services. (Customers employ a standard of zero tolerance.) Customers want to be treated well, never put down or demeaned. Customers also have come to expect an occasional value-added extra that makes it easier to do business with a company and improves the cost/benefit ratio. Finally, customers faced with a problem expect the supplier to recover, to fix the problem without harassing the customer.

Once customers have decided to purchase a product or service from a particular supplier, their overall satisfaction and willingness to do business with that supplier in the future rest with the supplier's ability to satisfy needs at the top of the hierarchy. Those who fail to manage the customer relationship at the top of the hierarchy loose customers despite the value, quality and availability of their products. Eventually, they create a reputation for themselves that waves off potential customers and erodes their sales base.

Any performance improvement effort should begin with an analysis of the company's performance against its customers' hierarchy of needs. Strengths and weaknesses should be identified and priorities set based on this analysis.

***Best Practice for Customer Satisfaction in Manufacturing Firms*, Griffin, Abbie, Greg Gleason, Rick Preiss, and Dave Shevenaugh, Sloan Management Review (Winter 1995)**

The most frequently measured Customer Satisfaction (CS) variables were expressed as numbers. Most companies use simple scales that assume satisfaction ranges linearly between 0 and 10 or 0 and 100. More elaborate measures of customer satisfaction that look at performance relative to expectations, or disconfirmation measures of satisfaction, are not frequently used. CS measures are often upwardly biased, not linear. Customers are the subset of the total population who are already somewhat satisfied with products and services, so the response population does not form a normal distribution about the midpoint, which is what most analytical procedures for linear scales assume. However, if you cut the scale off at 5 and consider only the responses above 5, the response distribution of the "average" firm might be much closer to a normal distribution about the new midpoint, 7.5 of 75 percent. This truncated scale would more closely conform to the standard statistical assumptions for linear interval scales.

The process of linking goals to performance through measuring CS is exploratory and preliminary at even the most forward-thinking companies. First, companies must formalize and quantify the relationship between CS and firm performance. By determining how CS improves performance or what specific CS components correlate with different improvements, corporations can focus on only the most effective endeavors, allowing them to become more efficient in implementation.

Delivering CS is at an early evolutionary state in most U.S. firms. Most firms are not focused on satisfying customers, even though research now correlates CS with improved performance. A firm's CS implementation process must reflect the needs of individual customer segments, and the overall program must be flexible enough to allow each business unit to develop measures and processes that fit its management needs.

***Avoid Top Box Problem by Using Another Box*, Dan Prince, President, Prince Marketing Research, Marketing News, June 1995, p. H-32.**

This article suggests an alternative to the "top box problem" when measuring customer satisfaction. This alternative uses a three-point scale. Respondents are asked to rate overall satisfaction, and satisfaction on individual attributes, as (1) much better than expected, (2) about as expected, and (3) worse than expected. If a customer chooses (1), it means they are expressing delight with the product or service, not just satisfaction. The research showed that if a customer is delighted, there is a 90% chance they will purchase the product or service again. If (2) is chosen, the customer is expressing satisfaction with a low product or brand loyalty. And finally, if (3) is chosen, the customer is dissatisfied with the product or service.

This alternative approach provides two benefits:

- it measures a customer's view against his or her expectation, and
- it gets rid of the top box problem of skewness — bias to the top of the scale.

Finally, using this alternative approach enables management to understand how well their product or service actually measures against their customers' expectations.

DEFINITIONS:

Top box problem:

> Most customers — if they are still your customer — will tend to give overall satisfaction scores that fall into one of the top boxes on your answer sheet, usually, "excellent" or "good" (7 to 10 on a 10-point scale).
>
> A second variation of the top box problem is that when respondents are asked, "How satisfied are you with X," followed with a request to rate X on a scale of importance, most customers will say each variable is either "very important" or "important."

***Rational and Adaptive Performance Expectation in A Customer Satisfaction Framework,* Johnson, Michael D., Eugene W. Anderson, and Claes Fornell, Journal of Consumer Research, Inc., Vol. 21, March 1995, pp. 595-707.**

There is an extensive and growing body of research on customer satisfaction that focuses primarily on disaggregate or individual-level satisfaction with particular goods or services. Relatively little attention has been paid to the determinants of market-level satisfaction, which is defined here as the aggregate satisfaction of those who purchase and consume a particular product offering (e.g., Ford Escort owners or Federal Express users). Studying customers in the aggregate is one way to establish empirical generalizations in the domain of satisfaction research.

The modeling of customer satisfaction depends critically on how satisfaction is conceptualized. Two general conceptualizations of satisfaction exist: transaction-specific satisfaction and cumulative satisfaction. Consumer researchers are often concerned with satisfaction as an individual, transaction-specific measure or evaluation of a particular product or service experience. Alternately, satisfaction is viewed as a cumulative, abstract construct that describes customers' total consumption experience with a product or service. This conceptualization of satisfaction is more consistent with existing views. Satisfaction is a customer's overall evaluation of his or her purchase and consumption experience to date. Measures of this satisfaction component can serve as a common denominator for describing differences across firms and industries, while transaction-specific evaluations provide information only about shortrun product or service encounters. Cumulative satisfaction is a fundamental indicator of a firm's (or market's) current and long-run performance.

To construct indices of customers' satisfaction at the market level for individuals who purchase and consume a particular product or service offerings, three measures are proposed: aggregate expectations, perceived performance, and satisfaction. Expectations are measured first by asking customers how well they expected the product or service to perform. Two measures are then collected to operationalize performance (perceived quality relative to price paid and a rating of how much the customer has paid

relative to how well the product or service has performed). Finally, three measures are used to operationalize satisfaction: overall satisfaction, confirmation of expectations, and distance from the customer's hypothetical ideal product or service in the industry. Three-stage (extrapolative, adaptive, and rational) least square estimates are used to determine market expectations and satisfaction. In every case, satisfaction is positively affected by both performance and expectations.

The results show that there is a significant carryover effect for customer satisfaction from period to period. That is, market satisfaction is a relatively stable, cumulative phenomenon that changes gradually over time.

Green, Paul E. and Tull, Donald S. Research for Marketing Decisions; 3rd edition; Prentice-Hall, Inc. 1975 (Englewood Cliffs, New Jersey), pp. 478-484.

In a typical customer satisfaction study, respondents evaluate overall satisfaction, followed by ratings on many individual attributes. A key question for researchers is which attributes are most important in determining overall satisfaction. Not all attributes have equal impact. A method of prioritizing is needed to allocate limited resources more efficiently.

Researchers have suggested many procedures for dealing with this problem. Several are considered by Green and Tull (1975), Hauser (1991), and *The Maritz Marketing Research Report* (1993). Work continues in this area; no true "answer" for all applications has emerged. However, derived importance measures are usually preferred over stated importance measures.

Stated importance measures ask respondents to explicitly state their perception of the importance of each attribute, usually using a 10-point scale. The results of this method can be straightforwardly interpreted; however, the results can be few, if any, statistical differences among attributes, so the aim of the method — to prioritize attributes — is thwarted. (How does a mean 7.8 rating differ specifically from a mean 7.5 rating?)

Derived importance methods rely on the statistical association between ratings (predictors) and an overall rating. The importance of an attribute is statistically determined from this relationship. Green and Tull consider four derived importance measures. If, in the very unlikely case that all attributes are uncorrelated with each other, all four yield identical measures of relative importance. Measures discussed by Green and Tull are:

- *Bivariate (Pearson) correlation:* This measure has the advantages of familiarity and simplicity. Unlike the other three, it's not affected by adding or deleting other attributes in a regression equation; however, joint effects with other attributes go undiscovered.

- *Standardized regression coefficient or beta weight:* Model misspecifications and the influence of other attributes in the regression model are particularly troublesome in this approach. This measure can be very unstable.

- *The product of the beta weight and the corresponding Pearson correlation:* This measure is a compromise between the two former measures.

- *The coefficient of part determination:* This model represents an incremental gain in predictive power but is adversely influenced by the inclusion or exclusion of particular attributes in the model.

All four measures exhibit limitations. However, an important consideration is that it is common in customer satisfaction research for attributes to be correlated — sometimes highly — with each other. This makes it difficult to measure the separate effects of the individual attributes on overall satisfaction. The latter three measures are all subject to instability when attributes are highly correlated. When interrelations exceed .5 — a fairly frequent occurrence for customer satisfaction data — the beta weights can shift dramatically.

Moreover, the latter three measures can also be affected by the addition or deletion of particular attributes to the regression model. The multiple regression model used for the latter three measures must have the correct functional form.

In the face of these problems, use of the first measure, simple bivariate correlation is recommended. However, considering each attribute in isolation is also unrealistic.

Green and Tull offer an alternative to combat multicolinearity; namely, to transform the original attributes into an uncorrelated set of new variables using the technique of principal component analysis. The principal components reveal the colinearity in the data while allowing analysis such as stepwise multiple regression to be performed without multicolinearity — and without deleting one of more of the highly correlated attributes.

This approach has the added advantage of using multivariate techniques that can be explained and described.

APPENDIX B

MODERATOR'S GUIDE

A. Introduction

This is a nationally based study to explore customer requirements for transit service. We want to know how riders view quality of service. What features of service are important? What are the most troublesome aspects of riding transit? How can a transit agency best improve its service? These are the kinds of questions we want to ask. We also want to know how you define quality service and get your reactions to various ideas about how a transit agency can monitor their quality of service. Let's start by having each of you introduce yourself.

1. Current transit usage, frequency of usage, trip purposes, how long have they been using transit, cars in the household, primary reasons for using transit over other modes of transportation.

B. Description of Ideal Transit Service

1. How would you define the 'ideal' transit service?

2. What would you change about your current transit service to make it closer to the 'ideal'?

3. How do you define low quality transit service?

C. Discussion of Basic Transit Requirements

1. What are the basic requirements for transit service?

2. How would you define the dimensions of service quality?

   — safety
   — comfort
   — ease of using the system
   — convenience
   — performance/reliability
   — facilities
   — value

D. Review of Specific Transportation Attributes

SAFETY

1. What does "safety" mean when using rail/bus?

2. Here are some features related to "safety" mentioned by others. How important is each in your decision to use transit?

   — Safety from crime while riding
   — Safety at stations/bus stops
   — Safety related to the behavior of other persons
   — Safety related to the rail/bus operation

3. Are there other aspects of "safety" we failed to discuss?

COMFORT

1. How do you define "comfort" when riding rail/bus?

2. Here are some features related to "comfort" mentioned by others. How important is each in your decision to use transit?

   — Availability of seating at the station/bus stop
   — Availability of seats on the train/bus
   — Smoothness of the train/bus ride
   — Comfort of the seats
   — Degree of crowding on the train/bus
   — Comfortable temperatures on the train/bus
   — Availability of handrails/grab bars

3. Are there other aspects of "comfort" we failed to discuss?

EASE OF USING THE SERVICE

1. How would you define an "easy" system to ride?

2. Here are some features related to "ease of using a service" mentioned by others. How important is each in your decision to use transit?

   — Knowing when trains/buses arrive and depart
   — Availability of information at a station (RAIL ONLY)
   — Availability of printed schedules
   — Ease of getting information by telephone

   — Courtesy/helpfulness of ticket agents (RAIL ONLY)
   — Ease of purchasing tickets/passes/tokens

   — Visibility of station names from on the train (RAIL ONLY)
   — Visibility of train/bus names/route numbers/colors from the outside

   — Ease of getting on/off train/bus
   — Ease of paying fare

   — Ease of making connections/transfers
   — Knowledgeable and courteous conductors/drivers on-board
   — Availability of information about delays from conductors/drivers
   — Clear/timely stop announcements

3. Are there other aspects of "ease of use" we failed to discuss?

CONVENIENCE

1. What does "convenience" mean when riding rail/bus?

2. Here are some features related to "convenience" mentioned by others. How important is each in your decision to use transit?

   — Availability of stations/bus stops close to home
   — Availability of stations/bus stops close to work
   — Availability of stations/bus stops close to shopping
   — Availability of parking at stations/bus stops

3. Are there other aspects of "convenience" we failed to discuss?

PERFORMANCE/RELIABILITY

1. What does "performance and reliability" have when riding rail/bus?

2. Here are some features related to "performance and reliability" mentioned by others. How important is each in your decision to use transit?

   — Frequency of service
   — Travel time by train/bus
   — On-time performance
   — Wait time when transferring

3. Are there other aspects of "performance and reliability" we failed to discuss?

CONDITION OF VEHICLES AND FACILITIES

1. How do you define vehicles and facilities in good condition?

2. Here are some features related to the condition of vehicles and facilities mentioned by others. How important is each in your decision to use transit?

   — Cleanliness of the train/bus interior
   — Trains/buses clean of graffiti
   — Stations/bus shelters clean of graffiti
   — Cleanliness of train stations/bus stops

3. Are there other aspects of the condition of vehicles and facilities we failed to discuss?

VALUE

1. How would you define "value" with respect to riding rail/bus?

2. Here are some features related to "value" mentioned by others. How important is each in your decision to use transit?

   — Cost of a one-way ride
   — Cost of a transfer
   — Availability of discounted fares, e.g., senior citizens, students
   — Availability of volume discounts, e.g., monthly passes
   — Cost of parking at stations/bus stops

3. Are there other aspects of "value" we failed to discuss?

E. Defining Service Quality

1. How should a transit agency measure/monitor its own quality?

2. What information should a transit agency collect and use to monitor its quality?

3. Reactions to _____ collecting the following quality measures.

   — percent of trips on-time
   — headway consistency
   — breakdowns
   — communication measures
   — number of accidents
   — vehicle availability

   If I told you that the _____ reports that 92% of all trips on the _____ line arrive within four minutes of their scheduled arrival time, what does that mean to you?

   What does it mean if I say that on average _____ buses break down every 3,500 miles?

4. Do these measures present an honest picture of the quality of service provided by _____?

5. How should a transit agency demonstrate that its customers come first?

F. Closing

1. What does quality of transit service mean to you as a rider?

BASIC DIMENSIONS

- ☐ safety
- ☐ comfort
- ☐ ease of using the system
- ☐ convenience
- ☐ performance/reliability
- ☐ facilities
- ☐ value

SAFETY

- ☐ Safety from crime while riding
- ☐ Safety at stations/bus stops
- ☐ Safety related to the behavior of other persons
- ☐ Safety related to the rail/bus operation

COMFORT

- ☐ Availability of seating at the station/bus stop
- ☐ Availability of seats on the train/bus
- ☐ Smoothness of the train/bus ride
- ☐ Comfort of the seats
- ☐ Degree of crowding on the train/bus
- ☐ Comfortable temperatures on the train/bus
- ☐ Availability of handrails/grab bars

CONVENIENCE

- ☐ Availability of stations/bus stops close to home
- ☐ Availability of stations/bus stops close to work/shopping
- ☐ Availability of parking at stations/bus stops

PERFORMANCE/RELIABILITY

- ☐ Frequency of service
- ☐ Travel time by train/bus
- ☐ On-time performance
- ☐ Wait time when transferring

EASE OF USING THE SERVICE

- ☐ Knowing when trains/buses arrive and depart
- ☐ Availability of information at a station
- ☐ Availability of printed schedules
- ☐ Ease of getting information by telephone

- ☐ Courtesy/helpfulness of ticket agents
- ☐ Ease of purchasing tickets/passes/tokens

- ☐ Visibility of station names from on the train
- ☐ Visibility of train/bus names/route numbers/colors from the outside

- ☐ Ease of getting on/off train/bus
- ☐ Ease of paying fare

- ☐ Ease of making connections/transfers
- ☐ Knowledgeable and courteous conductors/drivers on-board
- ☐ Availability of information about delays from conductors/drivers
- ☐ Clear/timely stop announcements

CONDITION OF VEHICLES AND FACILITIES

- ☐ Cleanliness of the train/bus interior
- ☐ Trains/buses clean of graffiti
- ☐ Stations/bus shelters clean of graffiti
- ☐ Cleanliness of train stations/bus stops

VALUE

- ☐ Cost of a one-way ride
- ☐ Cost of a transfer
- ☐ Availability of discounted fares, e.g., senior citizens, students
- ☐ Availability of volume discounts, e.g., monthly passes
- ☐ Cost of parking at stations/bus stops