The History and Future of Service Quality Assessment

Connecting customer needs and expectations to business processes

By Terry Grapentine

In recognition of the 10th anniversary of the publication of the SERVQUAL scale in the Journal of Retailing, the Academy of Marketing Science sponsored a special session on the advancements in service quality assessment over the last decade. This article reviews that special session.

There is a tide in the affairs of men, which, taken at the flood, leads onto fortune.” In the early 1980s, concerns about customer satisfaction and product quality became emerging tides in the affairs of industry and academia. And in the early 1980s, three academicians boarded their service quality boat and road the tide. And what a ride they had. Their work not only spawned numerous articles, books, conference presentations, and consulting engagements, but also significantly affected how many organizations went about measuring service quality.

In 1985, A. Parasuraman (pronounced PAH-RAH-SU-RAH-MUN, but he generally goes by his nickname, PAR-SU), Leonard Berry, and Valarie Zeithaml jointly published, “A Conceptual Model of Service Quality and Its Implications for Future Research,” which appeared in the fall issue of Journal of Marketing. Three years later, in the Journal of Retailing, they published their approach for defining and measuring service quality, SERVQUAL. Their unique contribution called for measuring both perceived performance (P) and customer expectations (E). One key to maximizing quality was to maximize the difference between these two measures, (P-E) -- in short, to exceed customer expectations.

No one can deny that those two articles and the Parasuraman, Berry, and Zeithaml troika had (and continue to have) a significant impact on the marketing research literature and industry. Because of that, the Academy of Marketing Science (AMS), at its 1998 Annual Conference on Current Developments in Marketing, sponsored a special session recognizing SERVQUAL’s 10th anniversary. The Academy of Marketing Science is a nonprofit organization whose mission is to “further the science of marketing throughout the world by promoting the conduct of research and the dissemination of research results through the study and improvement of marketing as an economic, ethical, and social force.” The Academy's official journal is the Journal of the Academy of Marketing Science.

Glenn B. Voss, North Carolina State University, organized the special session and invited, in addition to Parasuraman, one of the troika's major critics, R. Kenneth Teas, Iowa State University, and, as an independent observer and significant contributor to the service quality literature, Roland T. Rust, Vanderbilt University.
This article reviews each participant's presentation and provides a personal observation at the article's conclusion. In order to help the reader better understand the history and development of SERVQUAL, a more detailed discussion of its approach and methods is provided than what Parasuraman discussed at the conference.

SERVQUAL’S HISTORY

The Birth of SERVQUAL (1983-1985)

Parasuraman began the session by reviewing the history of SERVQUAL from the early 1980s to 1998. His collaboration with Berry and Zeithaml began in 1983. And prior to their seminal 1988 Journal of Retailing (JR) article, Parasuraman et al. published a conceptual paper in 1985 identifying five service quality gaps (see Exhibit 1). This framework is recognized today as a major contribution to the marketing literature, and it laid the foundation for what these three scholars were going to do over the next 13 years. Quoting from that article, the definitions of each of the gaps are as follows:

- **Gap 1**: Difference between consumer expectations and management perceptions of consumer expectations.

- **Gap 2**: Difference between management perceptions of consumer expectations and service quality specifications.

- **Gap 3**: Difference between service quality specifications and the service actually delivered.

- **Gap 4**: Difference between service delivery and what is communicated about the service to consumers.

- **Gap 5**: Difference between consumer expectations and perceptions.

SERVQUAL Instrumentation (1985-1988)

Of particular interest to Parasuraman et al. was Gap 5 -- the expected service/perceived service gap, which was the focus of the 1988 JR article. This article produced the famous equation, \( Q = P - E \), and operationalized it empirically. The operational definition of expectations \( E \) is given in Exhibit #2. The operational definition of \( P \) is given in Exhibit #3.

Exhibit 1

**Conceptual model of service quality**

```
CONSUMER

Word of Mouth Communications

Personal Needs

Past Experience

Expected Service

Gap 5

Perceived Service

MARKETER

Service Delivery (including pre- and post-contacts)

Gap 4

External Communications to Consumers

Translation of Perceptions into Service Quality Specs.

Gap 1

Management Perceptions of Consumer Expectations

Gap 2
```
The battery of items on which selected service providers were evaluated reflected the following service quality perceptual dimensions (quoting from the 1988 JR article):

- **Tangibles**: Physical facilities, equipment, and appearance of personnel.
- **Reliability**: Ability to perform the promised service dependably and accurately. **Responsiveness**: Willingness to help customers and provide prompt service.
- **Assurance**: Knowledge and courtesy of employees and their ability to inspire trust and confidence.
- **Empathy**: Caring, individualized attention the firm provides its customers,

The authors tested the model for one service firm in each of the following industries: banking, credit card, repair and maintenance, and long-distance telephone services. Model findings were statistically significant and the multi-item measures of the service quality perceptual dimensions demonstrated discriminate validity and possessed respectable coefficient alphas exceeding 0.70.

### The Extended Gaps Model (1988-1990)

Building upon their initial work, the authors offered an extended model of service quality, which is shown in Exhibit 4 (see page 8). In this extended model (Zeithaml et al. 1988), they identified a variety of factors, internal to an organization, that affect the level of service quality delivered to the customer. These internal factors are related to Gaps I through 4 in their original conceptual model.

**Gap 1: Difference Between Consumer Expectations and Management Perceptions of Consumer Expectations**

This gap is affected by the following three factors: (1) marketing research orientation of the organization, (2) upward communication, and (3) the number of levels of management in the organization.

**Marketing Research**: One medium through which management learns of consumer expectations is through marketing research. Consequently, the extent, nature, and quality of marketing research conducted by an organization can widen or lessen this gap. What's more, effective communication of marketing research can affect management perceptions of consumer expectations.
Upward Communication: "Though top managers may not have a firm grasp of consumer quality expectations, research suggests that customer contact personnel can accurately predict consumer expectations and perceptions of the service" (Zeithaml 1988, p.38). Thus, upward communication from customer contact personnel to top management can also affect Gap 1. Zeithaml et al. discuss at length the characteristics and nature of effective upward communication. For example, face-to-face communication is often more effective than written communication. "Management by walking around" and "employee gripe sessions" are examples of how the quality of upward communication can be enhanced.

Management Levels: Finally, the number of layers of management between customer contact people and top management will affect Gap 1. "Layers of management inhibit communication and understanding because they place barriers between senders and receivers of messages. Therefore, the greater the number of layers between customer contact personnel and top managers, the larger Gap I is expected to be."

Gap 2: Management Perception-Service Quality Specification Gap

The actual specifications management establishes for a service often differ from customer expectations. (Who hasn't been frustrated with cashing a check at one's local bank drive-up on a late Friday afternoon?) As shown in Exhibit 4, the size of Gap 2 can be affected by (1) management's commitment to service quality, (2) goal setting, (3) task standardization, and (4) perception of feasibility.

Management Commitment: If management is not truly committed to service quality, Gap 2 will suffer. Sometimes management's commitment is inconsistent. It only gets charged when customers (or an important customer) complain (or perhaps after a manager attends a customer satisfaction conference), and then, after a while, other concerns take precedence, such as short-term profitability and cost reduction objectives.

Goal Setting: "Companies that have been successful in delivering high service quality (e.g., American Express, McDonalds, Delta Airlines) are noted for establishing formal goals relating to service quality." Formal goal setting transforms pious platitudes into concrete actions and objectives. These, in turn, tend to standardize an organization's definition of what quality service is and

---

Exhibit 4

Extended model of service quality

- Marketing Research Orientation
- Upward Communication
- Levels of Management
- Management Commitment to Service Quality
- Goal Setting
- Task Standardization
- Perception of Feasibility
- Teamwork
- Employee-Job Fit
- Technology-Job Fit
- Perceived Control
- Supervisory Control Systems
- Role Conflict
- Role Ambiguity
- Horizontal Communication
- Propensity to Overpromise
- Gap 1
- Gap 2
- Gap 3
- Gap 4
- Gap 5 (Service Quality)
- Tangibles
- Reliability
- Responsiveness
- Assurance
- Empathy
how the organization will deliver quality service to the customer. 

**Task Standardization:** Standardizing tasks related to service delivery can significantly affect narrowing Gap 2. Think of how the fast food industry has standardized and routinized drive-up delivery service -- some arguably better than others. "According to Theodore Levitt, standardization or (in his terms) industrialization of service can take three forms: (1) substitution of hard technology for personal contact and human effort, (2) improvement in work methods (soft technology), or (3) combinations of these two methods." In short, the more that tasks can be standardized, the smaller Gap 2 will be.

**Perception of Feasibility:** Finally, Gap 2 is affected by the extent to which management believes it is feasible to meet customer expectations. The more feasible it is perceived to be, the more likely management will act. "Variables related to this construct include the organizational capabilities and systems for meeting specifications and the degree to which managers believe expectations can be met economically."

Clearly, the accuracy of information that management is given in regard to these issues will affect closing Gap 2.

**Gap 3: Service Quality Specification-Service Delivery Gap**

This is sometimes referred to as the service performance gap which occurs when employees are unable and/or unwilling to perform the service at the desired (i.e., management specified) level. Zeithaml et al. give the following examples of issues affecting Gap 3:

- **Teamwork:** Extent to which employees feel personally involved and committed.
- **Employee-job fit:** Ability of employees to perform job.
- **Technology-job fit:** Appropriateness of tools and technology for performing job.
- **Perceived control:** Extent to which employees perceive they are in control of their jobs.
- **Supervisory control systems:** Extent to which employees are evaluated on what they do (behaviors) rather than solely on output quantity.
- **Role conflict:** Perceived conflict between expectations of customers and expectations of the organization (e.g., a job that is too demanding).
- **Role ambiguity:** Perceived clarity of goals and expectations (e.g., employees do not have the information necessary to perform their jobs adequately).

**Gap 4: Difference Between Service Delivery and External Communications -- Piecrust Promises**

Horizontal communication and the propensity to overpromise can affect Gap 4. Horizontal communication refers to communication within and between departments in an organization. As an example, Zeithaml et al. cite an early Holiday Inn's "No Surprises" advertising campaign:
"Holiday Inn's agency used consumer research as the basis for a television campaign promising 'no surprises' to customers. Top managers accepted the campaign in spite of opposition by operations executives who knew that surprises frequently occur in a complex service organization. When the campaign was aired, it raised consumer expectations, gave dissatisfied customers additional grounds on which to vent frustrations, and had to be discontinued."

The "No Surprises" campaign is an example of what Mary Poppins calls a piecrust promise -- not easily made, but easily broken. This propensity to overpromise becomes greater the more a service firm feels pressure to generate new customers.

At the article's conclusion, the authors expound upon how this extended model can be tested empirically "by collecting data on the indicators of the five gaps through a cross-sectional study of service organizations and analyzing the data with a technique such as LISREL."

Nature and Determinants of Service Expectations (1990-1993)

During the early 1990s, Parasuraman, Berry, and Zeithaml continued to evolve their conceptual model. Their effort bore fruit in their 1993 Journal of the Academy of Marketing Science article, "The Nature and Determinants of Customer Expectations of Service."

This article's prime contribution is the development of the zone of tolerance concept (see Exhibit 5). Conceptually, the zone of tolerance is an area between a customer's adequate service level and the desired service level. For example, when you use a drive-up teller at a bank, your desired service level, in regard to waiting time, is most likely zero minutes. In order for the bank to keep you as a customer, you, on average, may not want to wait longer than 10 minutes -- the adequate service level. The difference between zero wait time and a 10-minute wait time is the zone of tolerance.

As Exhibit 5 shows, there are a number of factors that affect the zone of tolerance. For example, the zone of tolerance can be affected by situational factors ("It's payday and everyone is at the bank cashing a check, so I'll wait 15 minutes before I get upset"), past experience ("On Monday mornings I never have to wait in line, so I better not have to wait in line today"), and explicit service promises ("The bank advertises that I will never have to wait longer than 7 minutes at a drive-up teller").

This new model is based upon the following two propositions: (1) "Customers assess service performance based on two standards: what they desire and what they deem acceptable" and (2) "A zone of tolerance separates desired service from adequate service."


Parasuraman et al's. 1994 Journal of Retailing article tested alternative

---

**Exhibit 6**

The three-column format

We would like your impressions about _____'s service performance relative to your expectations. Please think about the two different levels of expectations defined below.

<table>
<thead>
<tr>
<th>Minimum Service Level</th>
<th>Desired Service Level</th>
<th>Perception of Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8 9</td>
<td>Low High</td>
<td>1 2 3 4 5 6 7 8 9 N</td>
</tr>
</tbody>
</table>

For each of the following statements, please indicate: (a) your minimum service level by circling one of the numbers in the first column; (b) your desired service level by circling one of the numbers in the second column; and (c) your perception of _____'s service by circling one of the numbers in the third column.
operational definitions of the zone of tolerance concept. It is beyond the scope of this review to discuss at length the nature of their empirical examination other than to show and briefly discuss the measure they found to be the most reliable and valid. Exhibit 6 (see page 10) gives their operational definition of the concept, which they refer to as the "three-column format."

The calculation of the zone of tolerance is straightforward -- subtract the minimum service rating from the desired service rating. The authors show the usefulness of their model in Exhibit 7.

Parasuraman et al. comment as follows:

Exhibit [7] "highlight[s] the suboptimal allocation of service-improvement resources that can result from using a perceptions-only measure to assess service quality, an approach advocated by some researchers. For instance, if the computer manufacturer examined just the perception scores it might decide to place the same emphasis on improving its performance on tangibles as on reliability. The imprudence of such a decision is evident from examining the company's performance on tangibles and reliability relative to the customers' tolerance zones for these dimensions. The data suggest that the company should place significantly greater emphasis on reliability than on tangibles."

Parenthetically, their observation does not take into account the relative importance of the dimensions. The priorities assigned by management to attribute performance improvement should take into account not only the tolerance zones for the attributes but also the relative importance of the attributes to customers. Parasuraman et al.'s 1994 JR article does not address the issue of the relative importance of the five dimensions, although some of their earlier publications do. For example, their 1988 JR piece reports importance weights derived through regression analysis, while their 1990 book -- Zeithaml, Parasuraman, and Berry, Delivering Quality Service: Balancing Customer Perceptions and Expectations -- describes and reports findings from a direct, constant-sum-scale approach for determining relative importance.

Impact of Service Quality on Behavioral Intentions (1994-1996)

In their 1996 Journal of Marketing article, "The Behavioral Consequences of Service Quality," Zeithaml, Berry, and Parasuraman reviewed the extant literature about the relationship between service quality and profits, and empirically examined several relationships between consumer behavioral intentions and service quality.

"Seminal studies using the PIMS (Profit Impact of Market Strategy) data

---

Exhibit 7

Service quality ratings for a computer manufacturer

<table>
<thead>
<tr>
<th>9</th>
<th>8</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reliability | Responsiveness | Assurance | Empathy | Tangible

---
set have uncovered significant associations among service quality, marketing variables, and profitability. Findings from these studies show that companies offering superior service achieve higher-than-normal market share growth, that the mechanisms by which service quality influences profits include increased market share and premium prices, and that businesses in the top quintile of relative service quality on average realize an 8% higher price than their competitors.

Building upon this literature, Zeithaml et al. developed a conceptual framework of the behavioral and financial consequences of service quality. Please refer to Exhibit 8. Superior (inferior) service quality is related to favorable (unfavorable) behavioral intentions. The authors examine the following behavioral intention measures (note that all but Internal Response are multi-item scales):

- **Loyalty**: The strength of the relationship between the customer and the service supplier and the likelihood the customer will do more business with the supplier in the future.
- **Switch**: The extent to which a customer might do less business with a supplier in the future and purchase from a competitor.
- **Pay More**: The likelihood a customer would be willing to continue purchasing from a supplier even though its prices increase somewhat.
- **External Response**: The likelihood a customer would complain to others and switch to a competitor if the customer experiences service problems.
- **Internal Response**: The propensity a customer would complain to a supplier’s employees if the customer experiences a problem with the supplier’s service.

Favorable (unfavorable) behavioral intentions are associated with customer retention (defection) which in turn is related to positive (negative) financial consequences.

Using their zone of tolerance model, Zeithaml et al. posit the following two hypotheses:

- **H1**: The service quality-behavioral intentions relationship (a) is positive (negative) for favorable (unfavorable) behavioral intentions and (b) has a different slope below and above the zone of tolerance relative to it.
- **H2**: Favorable (unfavorable) behavioral intentions are (a) highest (lowest) for customers experiencing no service problems; (b) next highest (lowest) for customers experiencing service problems that are resolved, and (c) lowest (highest) for customers experiencing service problems that are not resolved.

The empirical test of these hypotheses involved surveying business customers of a computer manufacturer and end customers of a retail chain, an automobile insurer, and a life insurer. Data were collected using a self-administered mail questionnaire.

Although the authors report some mixed results depending upon the
behavioral intention measure and the industry surveyed, their results by and large are encouraging. For example, when combining all industry data, "the quality-intentions relationship for loyalty and switch is flatter below, but remains unchanged above the zone [of tolerance]. Thus, exceeding the adequate-service threshold can sharply increase the payoffs (in terms of fostering customer loyalty and curtailing propensity to switch). However, the combined sample results for the pay more dimension reveal considerable flattening of the quality-intentions relationships above the zone of tolerance... Thus, companies wishing to improve service beyond the desired service level should do so cautiously and cost-effectively, because recouping the added expense by charging price premiums may not be a viable option."

Their data, as partially shown in Exhibit 9, support their second hypothesis. H2 predicts that customers experiencing no service problems will have the best behavioral-intention scores, followed by those experiencing problems that were resolved, and customers with unresolved service problems would have the worst scores.

In summary, Zeithaml et al.'s data support their claim that improving service quality will positively impact behavioral intentions. However, there may be diminishing returns. "Multiple findings suggest that companies wanting to improve service, especially beyond the desired service level, should do so in a cost effective manner." Customers may simply not be willing to pay for too much service quality;


Berry's and Parasuraman's 1997 article in the Sloan Management Review, "Listening to the Customer -- The Concept of a Service-Quality Information System," is the keystone that locks together all their conceptual and empirical work encouraging organizations to measure the quality of their customer service.

The Sloan article's thesis is that management must incorporate a dynamic service-quality information system that measures service quality from multiple perspectives and from multiple points of view. No single study at a particular point in time can capture the changing nature or the multiple facets of a service delivery system and its impact on the customer and the organization.

An effective service-quality

<table>
<thead>
<tr>
<th>Exhibit 9</th>
</tr>
</thead>
</table>

Mean behavioral-intentions scores classified according to service problem experience

<table>
<thead>
<tr>
<th>Behavioral-Intentions dimension</th>
<th>No service problem (Group 1; n=2153)</th>
<th>Service problems that were resolved (Group 2; n=455)</th>
<th>Service problems that were not resolved (Group 3; n=346)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loyalty</td>
<td>5.47</td>
<td>5.01</td>
<td>4.11</td>
</tr>
<tr>
<td>Switch</td>
<td>3.35</td>
<td>4.00</td>
<td>4.49</td>
</tr>
<tr>
<td>Pay more</td>
<td>3.76</td>
<td>3.63</td>
<td>3.11</td>
</tr>
<tr>
<td>External response</td>
<td>3.70</td>
<td>3.95</td>
<td>4.43</td>
</tr>
</tbody>
</table>

*The behavioral-intentions scores are on a 7-point likelihood scale.*
information system (SQIS) is one that periodically surveys customers, competitor customers, and employees. The following four research approaches can be used:

**Transactional surveys:** Service satisfaction surveys of customers following a service encounter.

**Customer complaint, comment, and inquiry capture:** Systems to retain, categorize, track, and distribute customer complaints and other communications with the company.

**Total market surveys:** Surveys that measure customers' overall assessment of a company's service. Research includes both external customers and competitors' customers, i.e., the total market.

**Employee surveys:** Surveys concerning the service employees provide and receive, and the quality of their work lives.

These four methods are essential components of any SQIS and should be used by virtually all types of organizations. Based on the type of organization and the products and services it offers, these methods can be supplemented with other approaches such as mystery shopping, focus groups, and customer advisory panels. And the frequency with which these projects are undertaken can vary, depending upon management's information needs and the nature of the market being served.

The authors outline the following five guidelines for developing a system that can ensure an effective SQIS:

**Measure service expectations:** Use the zone of tolerance concept to measure desired service, adequate service, and perceived service.

**Emphasize information quality:** The SQIS must generate information that is relevant, precise, useful, credible, understandable, and timely.

**Capture customers' words:** A service quality report showing that 4% of the customer base is very dissatisfied and another 13% is somewhat dissatisfied with the company's service may not get management's attention. However, if the report includes customers' verbatim comments, it may receive a very different reaction.

**Link service performance to business results:** This effort not only serves to motivate management to take action, it results in organizations developing "their own evidence of the profit impact of service quality to make the investment more credible and fact-based for the planning and budgeting process."

---

**Exhibit 10**

**Principal benefits of an effective SQIS**

- Encourages and enables management to incorporate the voice of the customer into decision making.
- Reveals customers' service priorities.
- Identifies service improvement priorities and guides resource-allocation decisions.
- Allows the tracking of company and competitor service performance over time.
- Discloses the impact of service quality initiatives and investments.
- Offers performance-based data to reward excellent service and correct poor service.

Reach every employee: Employees are the conduits through which service is delivered to customers. Employees must have information regarding what drives service quality and feedback on their performance in delivering quality service to the customer.

Exhibit 10 shows the principal benefits of a SQIS. Such a system helps to motivate an organization to provide quality customer service, instruct management and employees on how to achieve their quality service goals, and demonstrates the bottom line impact of a SQIS.

In summary, the Parasuraman troika has had a major impact on how organizations view and assess service quality. They have continually built upon their initial conceptual and operational definitions of service quality and have provided management with suggestions on how to integrate their ideas into service quality improvement programs. Many would argue that they have made a lasting impression on their industry.

A CHINK IN THE SERVQUAL ARMOUR?

Rumblings of Au contraire.

That was Teas' response to the initial conceptual and operational definitions of expectations in his 1993 JM article, "Expectations, Performance Evaluation, and Consumer's Perceptions of Quality." (Parenthetically, I need to point out that, notwithstanding Teas' critique of SERVQUAL, he does recognize the contribution Parasuraman et al. have made to the literature. His observations are made in the spirit of building upon and improving their conceptual and operational models). In that and subsequent articles, Teas developed a number of arguments attacking SERVQUAL both theoretically and empirically. His discussion at the conference, however, focused primarily on theoretical and conceptual issues. To better understand the argument Teas made at the conference, one needs to briefly review his article.

Exhibit 11 shows an earlier operational definition of expectations (Parasuraman et al. 1990). Teas asks the question, "What does it mean, for example, if a respondent does not circle a 7 rating for the statement, 'Excellent ______ have up-to-date equipment?'") Such a rating implies that a respondent does not deem it essential for an excellent company to possess this feature.

But what do the words essential for excellent companies mean? Teas explored this issue by administering this expectations measure to a random sample of respondents. Respondents who did not circle a 7 for a particular SERVQUAL attribute were asked why they did not do so. The findings in Exhibit 12 show that a considerable proportion of the 432 responses to the question reflect that many respondents...
interpreted the expectation question differently. In other words, the phrase, *essential for excellent companies*, in the question of Exhibit 11 is ambiguous and introduces measurement error in how respondents answer the question. For example, several respondents did not circle a 7 because it is not feasible for an excellent company to perform at that level, whereas other respondents did not circle a 7 because the attribute is not important to them.

As Teas articulated at the conference, the foundation of his argument is that *any* operational definition and any empirical support of a model must rest on strong, theoretical footings. If not, these definitions become vague, ambiguous, and ultimately affect the validity of the quantitative models on which they are based.

What makes Teas’ argument somewhat controversial in some quarters is that he believes there are few conceptual definitions in the entire field of marketing that rest on solid theoretical ground. And a primary cause of this is the lack of the use of a formal language system in our field. As a result, one is faced with the formidable task of using natural language to express precise theories of consumer behavior and to differentiate one’s ideas from others. Teas, Shelby D. Hunt of Texas Tech University, and others contend that natural language often has insufficient precision to adequately express theory. For example, in his 1997 *JM* article, Teas documented 19 different definitions of expectations that are not clearly differentiated from one another.

### Vagueness and Ambiguity

**Associated with the Desired Concept**

Teas points out that *vagueness* and *ambiguity* are different terms. Ambiguity means that a term or concept can possess multiple meanings. A term or concept is vague to the extent that its meaning is not clearly defined.

Focusing primarily on Parasuraman et al.’s later work, Teas quoted the SERVQUAL authors as stating that *desired service* is "the level of service representing a blend of what customers believe 'can be' and 'should be' provided" (PZB, *(JR)* 1994, p. 202). However, many of these definitional terms are as vague as the terms used in an earlier expectations measure (e.g., "essential

### Exhibit 12

**Reasons for non-extreme ("non-7") expectations scale responses**

<table>
<thead>
<tr>
<th>SERVQUAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
</tr>
<tr>
<td>Not feasible</td>
</tr>
<tr>
<td>Sufficient</td>
</tr>
<tr>
<td>Not feasible/sufficient</td>
</tr>
<tr>
<td>Not feasible/importance</td>
</tr>
<tr>
<td>Not feasible/ideal</td>
</tr>
<tr>
<td>Sufficient/important</td>
</tr>
<tr>
<td>Ideal</td>
</tr>
<tr>
<td>Ideal/importance</td>
</tr>
<tr>
<td>Forecast</td>
</tr>
<tr>
<td>Forecast/importance</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Total &quot;non-7&quot; responses</td>
</tr>
</tbody>
</table>

1In this table the "Not necessary" and "Sufficient" categories are combined.

2In this table the "Less important" and "Implied less important" categories are combined.

for excellent companies").

What does "can be" mean? Does "can be" mean what is feasible for a company to provide; does it refer to some evaluation of what a company can provide based on past experience; or,
perhaps, does "can be" refer to an ideal level of performance?

What does "should be" mean? Does "should be" refer to what a customer feels is equitable, based on what the product costs, or does this term reflect a forecast of future performance, based on past performance? If respondents have different notions as to what these concepts mean as they are related in survey instruments, their answers to survey questions will contain measurement error (as demonstrated in Exhibit 11). Data variance will partially be due to how respondents interpret the question, not to true differences in respondent attitudes or perceptions. Additionally, to the extent that marketing scholars cannot effectively communicate with each other because of the inherit shortcomings of natural language, confusion is introduced into the literature.

Vagueness and Ambiguity Associated with the Adequate Service Concept

Teas quoted Parasuraman et al. as stating that adequate service is "the minimum level of service customers are willing to accept" (ZBP, 1994, p. 203). Examples of vague terms in this definition are "minimum level of service" and "willing to accept." What exactly happens at this presumably low service level? Do consumers switch brands? Do they grudgingly accept the low service level hoping it will improve? One could argue that these two consumer actions could occur at two different levels of service. If the definition of these terms is unclear to respondents, measurement error is introduced into their answers.

Vagueness and Ambiguity Associated with the "Service Quality" Concept

Teas cites the following quotes regarding the meaning of service quality:

- "...Perceived service quality is viewed as the degree and direction of discrepancy between customer perceptions and desires (PZB, 1985, 1986, and Zeithaml et al. 1993).
- "...Assessments of service quality result from a comparison of desired service with perceived service" (Zeithaml et al. 1993).
- "...The comparison between desired service and perceived service, which we call perceived service quality (PSQ) Gap 5a, is the perceived service superiority gap..." (Zeithaml et al. 1993).

Following Teas' reasoning, assume the letter P denotes a perceived service measure on a 1- to 7-point scale where higher numbers denote better service. The letter D denotes the desired service level and the letters SS denote a measure of superior service, which is defined as P - D. If P exceeds D, then a company is providing superior service, according to Parasuraman et al. Consider the logic of the following relationships:

- P(l) = 7 and D(l) = 6, therefore, SS(l) = +1
- P(2) = 6 and D(2) = 5, therefore, SS(2) = +1
- Therefore, SS(l) = SS(2), even though performance P(l) is higher than performance P(2).
According to Teas, this is logically inconsistent. And this issue was explored in the following discussion among Teas, Parasuraman, and myself:

Grapentine: How do you address Teas' example where SS(1) is equal to SS(2), even though P(1) is higher than P(2)?

Parasuraman: The same kind of gap score can be derived from different combinations of expectations and perceptions scores, so they all imply the same thing. This issue has come up before. And it is not an issue that can be resolved simply. But if you look at it [from] an individual perspective; if you look at it devoid of the individual that is making the evaluations, then you'd say, ok, the +1 [score of 7-6 = + 1] equals +1 [6-5 = +1]. But when you look at the specific individual, there is almost a common method bias here. Maybe the particular individual is a hard rater, so [he/she doesn't] give high values on any kind of rating scale. So from that perspective, I think one can argue that even though the second person's score is only a 6 as opposed to a 7 for the first person, they might still be experiencing the same level of gap, which is really what we are trying to measure. But the differences in the cross scores that leads to the gap are imperfect because of scale response biases or tendencies [to answer scales] that different individuals might have.

Grapentine: What if this is aggregate data and we are averaging over, say, 200 people. Doesn't your model suggest that A and B are equal levels of quality?

Parasuraman: It is highly unusual that you'd get that in the aggregate, is one response that I would give. I think at the individual level this can happen, but at the aggregate level, unless there are significant changes in the service itself, or in the environment itself, this is not going to happen.

Teas: I want to make a couple of points. When you deal with the question you asked, there are two domains [we are dealing with]. One is the measurement validity question, which is what [Parasuraman] addressed, and [what he said] could very well be the case. In fact there are studies that show exactly what [he] is talking about. My perspective, because of my assigned topics [in this special session] would be a little different than that. I believe that when we construct a theory, we have to keep separate [1] measurement artifact explanations for what happens vs. [2] theoretical implications. A theory has to be internally consistent, and independent of measurement error. And that's where I have a little bit of a problem with [his explanation] because I want the theory to hold independent of measurement error. And in that case, the measurement error. would not be an explanation. It may be as simple as recognizing that we need to model in a direct linkage of performance independent of the gap, like what has been done in the disconfirmed expectations theory.

Parasuraman: I'd also add that in our own work, partly in response to the debate that has occurred, we no longer look at just these gaps. If you have the zone of tolerances up there, then you're getting a much better picture. Look at the full picture, not just the gaps. I think you are much less likely to be misled. In a follow-up e-mail correspondence to me, Parasuraman elaborated on this point further:

"To the extent that this individual's desired level of service for [attribute] B is one scale point less than for [attribute] A, the 'service superiority' experienced by him/her on B will/can logically be the same as that experienced by him/her on A. [This can occur] even though the perceived performance on A is one scale point higher than on B. Of course,
implied in this argument is the assumption of linearity i.e., equal increments in performance relative to the desired level produce equal increments in the individual's 'psychological state' (of pleasure/displeasure).

One other point Teas makes is in regard to the desired service level, D. "Why," Teas asks, "wouldn't a rational respondent always give the highest rating for D on any attribute?" Teas says that the final version of the "three-column format" for measuring the desired service level (Parasuraman et al. 1994) is as follows: "DESIRED SERVICE LEVEL -- the level of service performance you desire." We have to assume that respondents evaluate these attributes independently from each other. Otherwise, we confound our ability to interpret the attribute ratings. Furthermore, respondents should not be making assumptions on how their desired service level would affect product cost. So, why, in Parasuraman et al.'s data do we observe non-optimal ratings on the D scores? This is further evidence, Teas states, that their D measure is not measuring what it purports to measure.

An alternative wording for the desired service level is as follows: "...the level of service that can be and should be provided." This alternative wording may explain less than 7 ratings; however, Teas questions the potential ambiguity of the terms, can be and should be.

The Elasticity of Loyalty with Respect to Service Quality

Teas asks how one interprets the change in elasticity of loyalty with respect to service quality, when perceived service quality exceeds the zone of tolerance, as shown in Exhibit 13. Teas says there are several factors that could explain the diminished slope of the Loyalty function beyond the zone of tolerance in Exhibit 13. One reason may be a respondent's reluctance to use the largest scale ratings (a situation Parasuraman mentions in the previously quoted dialogue). Thus, the reduced slope of the Loyalty function may be due to how respondents use the scale, not to a true change in the elasticity of loyalty with respect to service quality.

The Zone of Tolerance and Attribute Importance

Referring to Exhibit 7, Teas asks the following questions: "If one knows the relative importance of each of the attributes, how does knowing the zone of tolerance for each attribute contribute to making better decisions?" For example, let's assume that some reliable method is used to measure attribute importance and that the tangibles attribute is the most important attribute and the reliability attribute the least important. Then management clearly would place more importance on improving the tangibles attribute than the reliability attribute, regardless of the information contained in the zone of tolerance construct.

In summary, Teas' efforts have been channeled toward building upon and improving Parasuraman et al.'s conceptual frameworks and operational methods. His primary message is that to fully realize the potential of their ideas, they and others who use or plan to build upon their work, need to improve the preciseness of the SERVQUAL conceptual and operational definitions. Doing so will not only further the pursuit of truth but also make these service quality models more useful and valuable to managers.

"CUDGEL THY BRAINS NO MORE ON IT," RUST EXCLAIMS!

In his contribution to the session, Roland Rust outlined nine key personal observations and future research areas based on (1) Parasuraman's and Teas' academic work, and (2) Rust's own experiences in the "real world" as a marketing and marketing research consultant.
He prefaced his presentation by stating his admiration of Parasuraman et al.'s body of work in the area of service quality assessment as well as Teas' contribution to the literature. Nevertheless, Rust's mission at this session was not to lavish his approbation on these scholars, but rather to entertain the audience and to provoke our thinking on other issues that are at least equally important in assessing service quality. The general thrust of Rust's observations was that academics and practitioners should allocate more of their attention to acting on their research findings and less on examining the nuances of measuring customer satisfaction. His nine observations are discussed below.

**Item #1: Don't Sweat the Small Stuff.**

There is a certain depth to which academics can go in their investigations of measurement and marketing theory that can produce valuable and useful information to marketing and marketing research practitioners. But beyond that point, their findings may have little practical value.

"Here's some stuff that is small stuff that we should stop worrying quite so much about: (1) Number of scale points, (2) type of scale, and (3) is customer satisfaction the same as or different than service quality and vice versa? First of all, the number of scale points -- three, five, seven or eleven. I don't care. I've used all these lengths of scales and I don't think it makes a whole lot of difference. Other issues that are involved in service quality measurement swamp the difference that it makes. So let's stop worrying about it. Second of all, [with regard to the] type of scale [used], I've seen different companies use different scales -- disconfirmation scales, service quality scales, a customer satisfaction scale, or they are using a bunch of these scales and putting them together -- and I think they get more or less the same [result]. Now it's not exactly the same thing, but not enough different to matter. Third, if you keep cumulative and transaction constructs distinct, service quality always leads to customer satisfaction in the literature, and that is very consistent across all the literature. Finally, there are many customer satisfaction psychologists out there looking at little tiny things in terms of how the psychology of customer satisfaction works. I think that is great, but I don't think business cares one hoot about that right now. What they care about is some bigger issues [such as linking customer satisfaction to financial performance and producing a clearer managerial decision focus]."

**Item #2: Return on Quality.**

Presumably, the depths to which Parasuraman and Teas have taken their investigations pales in light of management's need to quantify and realize a return on their investment in quality. In short, business wants to know, "What is the return on quality?"

In this vein, Rust points to management's need to better utilize research results. How can management translate research findings into making more profitable decisions? And how can management use research as a strategic lever and better decide what it really needs to change to make a real difference in the company's bottom line?

It was at this point in Rust's presentation that he and I engaged in the following dialogue:

**Grapentine:** Roland, if you don't sweat the small stuff, how do you know the decisions [management makes on research findings] are correct?

**Rust:** I think that the small stuff doesn't make that much difference, because you can look at [a research issue] two or three different ways and you get the same decision every time. So whether you have a satisfaction or a
service quality scale, for example, it doesn't seem to matter. You end up weighting the same things the most. And my research has demonstrated that over time and other people's research has tended to demonstrate that. Practitioners will say, almost unanimously, that they looked [at an issue] several different ways and it doesn't seem to matter, which one they do. That's what you here from practitioners.

Rust's statement has profound implications for our field, if he is correct. Parenthetically, I would like him to document this claim and submit it to a referred journal.

**Item #3: Customer Delight.** Rust contends that management does not want simply to satisfy customers, but rather to delight them. And this is an important point because the actions management takes are different in regard to satisfying customers vs. making them delighted. "If you are merely trying to satisfy somebody, you take problems away and make sure the problems are solved. If you are trying to delight somebody, that's adding extras. Those are two different managerial approaches [to the customer] and that's why it becomes important."

Researchers, therefore, should avoid sweating the small stuff and direct their attention to discovering those actions that produce delight and those that don't.

**Item #4: The Bridge to Actionability.** Academics and practitioners focus too much on measuring customer satisfaction or expectations and too little on connecting customer needs to business processes. "'Me key question here is how do we implement these changes within the organization and how do we set up our measurement such that it connects well to the organization. How do we manage the interface between the measurement process and the organizational change process?"

**Item #5: CSM as a Management Philosophy.** Rust contends that the very act of conducting customer satisfaction research has deep philosophic implications for management that truly believe in CSM and isn't merely giving it lip service. First, it implies an outside-in view of the organization. The company must listen to what the customer wants them to do. Second, a commitment to CSM is also a commitment to revenue and growth "quotas" based on satisfying and delighting the customer. Finally, a true commitment to CSM suggests a new role for the marketing function. Rust explains as follows:

"If you open a typical marketing or marketing management book, you see the four Ps as a guiding structure. But what are the four Ps? It's a bunch of decisions you make about products. That's inside-out [thinking]. That is a product focus. But what we should be doing is making decisions about customers. Customers is where it's at, not products. We shouldn't be talking about brand equity, we should be talking about customer equity. That means marketing is about customer satisfaction and about the lifetime value of a customer. Marketing is about the customer first and that a company is organized around a customer, not organized around a product. We shouldn't have four Ps oriented books. I think that is inappropriate for a principles of marketing course."

As you can see, Rust is a bit of an iconoclast, which, perhaps, explains why
he was invited as the independent third party observer at this session.

**Item #6: A New Twist on Quality.**
The quality construct that came out of manufacturing is obsolete in today's marketing environment. The manufacturing focus on quality is myopic and tends to direct management's attention to "zero-defects" issues and manufacturing productivity. "When you start talking about services," Rust contends, "suddenly you start talking about trade-offs between customer satisfaction and productivity. You want high customer satisfaction, you are going to have to give up some productivity. You want high productivity and efficiency, you are going to have to give up some customer satisfaction."

**Item #7: Where Does Market Share Come From?** The conventional wisdom says that market share comes from customer choice. Every time a customer makes a product decision, market share is created.

That's the old way of looking at the customer. That doesn't work in today's world. Sure, when a customer first purchases a product, an initial choice is made. But after that, management is better off focusing on customer retention issues as opposed to merely the next purchase.

Management needs to separate initial customer choice issues from customer retention issues and discover what drives each. Separate models should be constructed for understanding how the initial customer choice decision is made and what it takes to keep customers after they make their first purchase.

**Item #8: Multicollinearity.** "This," Rust says, "drives practitioners crazy. This is because the coefficients [that are derived by statistical models such as multiple-regression analysis] are unstable." Rust noted that this is a research area in which he is currently involved, and it is an example of a future research area that is more important than deciphering whether five- or seven-point scales are better, or debating what comes first, customer satisfaction or service quality.

**Item #9: How Are Expectations Updated.** Notwithstanding the Parasuraman/Teas debate on defining "expectations," both conceptually and operationally, Rust's final point in his presentation focused on the need for future research to investigate how consumer expectations are updated -- that is, what is the process by which consumers revise their expectations of product performance. He briefly discussed three related subtopics -- linear updating, Bayesian updating, and threshold effects -- and how this is an important subject for future research.

In conclusion, Rust's main point is that the "small issues have been beaten to death" in the academic literature. Further articles on the ideal number of scale points, splitting hairs on defining expectations, and exploring the minutiae of customer psychology as it relates to customer satisfaction, may contribute to a professor's receiving tenure, but such research contributes little to solving the more important issues confronting today's management. What's more, "fine tuning" the answers to the "small stuff" isn't going to make a difference in research findings. As long as researchers are using generally accepted (although not necessarily universally accepted) research procedures, research findings will be about the same.

Many bigger issues need more work. For example, researchers should divert a significant amount of their effort in studying the "small issues" to exploring more important and actionable issues such as understanding the linkages between CSM and a company's financial performance. Additionally, research efforts need to focus on identifying factors that cause customer satisfaction and delight and understanding how CSM can best be utilized to affect significant changes in an organization.
To this end, Rust calls for an expanded dialogue and working relationship between academics and practitioners. "Academics need to listen more to the real world," Rust states, "and practitioners need to listen more to academics."

**A PERSONAL OBSERVATION**

This special session on service quality assessment reflects a natural evolution in the field of marketing and marketing research. This evolution is the growing role of theory in our discipline and the importance of definitions in our theories. Concomitant with this evolution are the practical demands of our field, as Roland Rust reminds us. Academics must translate their theoretical and empirical investigations into a language that practitioners can understand. Their findings ultimately must help management make more effective decisions. And management must be given the tools to know they have made the correct decisions.

All scientific disciplines begin empirically. We observe the world around us and try to explain what we see. And our first explanations are communicated in our native languages (i.e., natural language systems) as opposed to a formal language system.

Even though all of the "hard" sciences began in this way, natural language systems inhibit scientists' abilities to think clearly, precisely, and to ask the right questions about the phenomena they observe. Theory and sound theoretical definitions must guide empirical investigations. The field of marketing research needs to develop and agree upon a formal language system to define its conceptual definitions; otherwise, debates such as the one between Parasuraman and Teas will be never be resolved.

This is particularly important in our field. Marketing is not a hard science. Unlike many sciences such as biology, chemistry, and physics which can objectively observe much (but not all) of what they study, marketers cannot observe consumer attitudes and perceptions. In fact, marketing constructs and terms such as "attitude," "perception," "satisfaction," and "loyalty," do not have objective reality. They cannot be observed by marketing researchers who want to measure them or by consumers whose minds create them. The reason a construct is called a construct is because marketing scholars and practitioners construct them. We make them up. They are ideas we use to try to understand how humans think about things and how these thoughts affect behavior. Often, we make the mistake and think they really exist -- a problem called reification. Because of this, good definitions in our field are all the more important in our quest to understand the consumer.

So, are Parasuraman et al. and R. Kenneth Teas sweating the small stuff, as Roland Rust declared at the beginning of his presentation? I don't think so. And as impractical and "theoretical" as their dialogue may appear to be ostensibly, when you look below the surface, there are few things as practical and potentially useful to marketing practitioners as their past, present, and future dialogues.

The Academy of Marketing Science's special session, Reflections on Service Quality Assessment: Past, Present and Future, reflects the dynamic, and we all hope, progressive evolution of our field.

And for our field to progress we need people like Parasuraman, Berry, and Zeithaml to help us look at marketing issues from a new, fresh, and useful perspective. Their conceptual models have given management all of this. They truly have changed the landscape of the service quality literature and the way academics, research practitioners, and managers think about quality. Our field is better off because of their work.

We need people like Teas to help us think more clearly, more precisely, and to ask the right questions about...
marketing issues. He has begun the development of a formal language system for the field of marketing research, which, if used and built upon, will change the way we define, discuss, and measure marketing constructs.

And we need people like Rust to remind us that there are other important issues in addition to measurement issues that demand the attention of academics and practitioners. Academics and practitioners need to work together to demonstrate to management how to establish service quality measurement systems that improve organizational efficiency. And they need to work together to better enable management to implement service quality changes within the organization more effectively.

**ADDITIONAL READING**


