Pragmatic Perspectives on the Measurement of Information Systems Service Quality

By: William J. Kettinger
Center for Information Management and Technology Research
Department of Management Science
College of Business Administration
University of South Carolina
Columbia, SC 29208
U.S.A.
bill@sc.edu

Choong C. Lee
Franklin P. Perdue School of Business
Information & Decision Sciences
Salisbury State University
Salisbury, MD 21801
U.S.A.
cclee@ssu.edu

Abstract

In this research note, we join the debate between Van Dyke, Kappelman, and Prybutok and Pitt, Watson, and Kavan pertaining to the conceptual and empirical relevance of SERVQUAL as a measure of IS service quality. Adopting arguments from marketing, Van Dyke et al. (1997) question the SERVQUAL gap measurement approach, the interpretation and operationalization of the SERVQUAL expectation construct, and the reliability and validity of SERVQUAL dimensionality. In a response to those arguments, Pitt et al. (1997) defend their previous work (1995) in a point-by-point counterargument that suggests that the marginal empirical benefit of a perceptual-based (SERVPERF) service quality measure does not justify the loss of managerial diagnostic capabilities found in a gap measure. While siding with many of the positions taken by Pitt et al. (1997), we attempt to add value to the debate by presenting discrepancies we have with the two other research teams and by suggesting alternative approaches to resolve, or at least alleviate, problems associated with SERVQUAL. We believe that the theoretical superiority of an alternative IS service quality measure should be backed by empirical evidence in the IS context, hence answering some of the criticism by Van Dyke et al. and offering a construct valid version of the IS-adapted SERVQUAL. From a pragmatic viewpoint, we believe that the justification of using SERVQUAL's gap measure should be driven by more effective ways to utilize expectations in IS service management. To this end, we introduce the newer Parasuraman et al. (1994b) measures, the concept of a "zone of tolerance" for expectation management and an illustration of its practical use in an IS setting. Overall, we attempt to set the direction of where we think this debate should lead the IS field, namely, toward practical and timely IS service quality measures.

Keywords: IS service quality, measurement, SERVQUAL, SERVPERF, IS management, evaluation, user expectations

ISRL Categories: AI0104, AI0109, A104, E10206.03, GB02, GB07

Introduction

SERVQUAL is widely used today as a diagnostic tool for uncovering areas of service quality strengths and shortfalls. The practical value of SERVQUAL is twofold. First, SERVQUAL can be used as a benchmarking
tool. Benchmarks, to establish service “best practice,” can be drawn by comparing the summary SERVQUAL scores of major “players” within the same industry. Second, SERVQUAL can be used as a diagnostic or prescriptive tool. In this way, periodic measures of dimensional scores can identify problems within specific service processes. SERVQUAL’s importance, particularly as a diagnostic tool, is propelling researchers to refine the measure.

In developing their SERVQUAL instrument, Parasuraman, Zeithaml and Berry intended to derive a service quality measure that would transcend multiple measurement contexts. The original 22-item SERVQUAL (Parasuraman et al. 1988) instrument was based on extensive focus group research (Parasuraman et al. 1985) and designed for application in a broad spectrum of service sectors, including an intra-firm context (Zeithaml et al. 1990). On the basis of multiple empirical studies (Parasuraman et al. 1988, 1991, 1994b), SERVQUAL’s five dimensions include: (1) Tangibles: The appearance of physical facilities, equipment, personnel, and communication materials; (2) Reliability: The ability to perform a promised service dependably and accurately; (3) Responsiveness: The willingness to help customers and to provide prompt service; (4) Assurance: The knowledge and courtesy of employees and their ability to convey trust and confidence; and (5) Empathy: The caring attitude which provides individualized attention to customers. They assert that as core evaluation criteria, SERVQUAL provides a basic “skeleton” underlying service quality that can undergo minor wording modifications and be supplemented with context-specific items to fit specific needs (Parasuraman et al. 1988, 1991).

Recognizing the need to more comprehensively measure information services (IS) quality, Kettinger and Lee (1994) adapted the SERVQUAL instrument to the IS context as an enhancement to the existing User Information Satisfaction (UIS) measure (Ives et al. 1983). In their article, Kettinger and Lee (1994) discussed SERVQUAL’s theoretical roots in the quality and consumer satisfaction literature and focused on the practical value of the instrument’s flexibility in the IS service context. Pitt et al. (1995) further extended the application of SERVQUAL in IS by placing service quality within the IS Success Model (Delone and McLean 1992) and by independently testing SERVQUAL’s reliability and validity in samples from three different organizations. The strength of the IS-adapted SERVQUAL instrument was then examined cross-culturally, using organizations in four different countries (Kettinger et al. 1995). In these introductory papers, the conceptual emphasis placed service quality within the realm of past IS evaluation research and the empirical focus was primarily on determinations of IS-adapted SERVQUAL dimensionality. These articles only briefly touched on questions related to the use of “gap” scores and potential problems in the SERVQUAL “expectation” measure.

While these SERVQUAL papers were working their way through the IS review process in the early 1990s, a vigorous debate was taking place in the marketing field concerning the conceptual and empirical integrity of the SERVQUAL instrument. A research note appearing in this issue of MIS Quarterly (Van Dyke et al. 1997), provides a thoughtful critique on the IS-adapted SERVQUAL. The primary focus of this critique is the difference score conceptualization (i.e., perceptions minus expectations) invoked by SERVQUAL to operationalize service quality. Drawing heavily from the previously mentioned marketing debate, the critique points to conceptual and empirical difficulties with the original SERVQUAL instrument that have been cited in the past and suggests that alternatives to the original “gap scored” IS-adapted SERVQUAL should be used.

In a well constructed response, also appearing in this issue, Pitt et al. attempt to paint the remainder of the IS service quality “canvas” by clearly articulating a point-by-point discussion balancing many of the positions of Van Dyke et al. In defending their work, they offer counterarguments from the marketing literature. They concede that perception-only measures of service quality may have marginally higher convergent and predictive validities than SERVQUAL difference scores but they believe...
the loss of managerial diagnostic capabilities calls for continued use of the gap measure.

By reading both of these articles, one gains a thumbnail view of a debate that has been waged in marketing over the past four years. As for our research, while we appreciate the important contributions of Van Dyke et al., we generally side with many of the positions of Pitt et al. and are resistant to "throw the baby out with the bath water" until strong conceptual, empirical, and practical evidence pushes the scale markedly away from SERVQUAL. This stance is taken for several reasons. First, the debate in marketing over SERVQUAL is ongoing and, as will be discussed, researchers have moved to address many of the problems cited by Van Dyke et al. Second, the IS-adapted SERVQUAL instrument has been the subject of a series of empirical works within the IS field and we believe that the theoretical superiority of an alternative IS service quality measure should be backed by empirical evidence in the IS context. Third, considering the urgent need for service quality measurement in practice, the IS field must weigh the practical value of continually refining a useful, but less than perfect, service quality measure in lieu of having no tested alternative. Our pragmatic stance on this matter will be presented in response to the two main criticisms of Van Dyke et al.: (1) problems with SERVQUAL's dimensionality and (2) conceptual and empirical difficulties with the SERVQUAL gap measurement.

As both Van Dyke et al. and Pitt et al. have done a fine job in presenting the marketing arguments from both sides of the fence, we will try to avoid reiterating these positions in detail. Instead our response focuses on their representation of our work and other issues where we can add value to the debate. As SERVQUAL is so closely tied to practice, it is also our aim to concentrate on the implications of this debate for IS managers and new and potentially more fruitful avenues of research aimed at improving measures of service quality.

**SERVQUAL Dimensionality: Insurmountable Hurdles?**

The SERVQUAL instrument is designed to provide managers with deeper insights concerning the dimensions of IS service quality. Knowledge of these dimensions can provide practitioners with potentially useful diagnostics. For example, analysis of SERVQUAL dimensionality may direct an IS manager to deploy more staff toward improving the reliability of the system maintenance process rather than dedicating more staff toward improving responsiveness to customers through additional one-on-one help desk sessions. Used periodically to benchmark across IS business processes, SERVQUAL's total and dimensional scores have the potential to direct continual IS service quality improvement. However, Van Dyke et al. report that results from both Kettinger and Lee (1994) and Pitt et al. (1995) empirically demonstrate poor discriminant validity, diminishing the value of such diagnostics.

We agree that the dimensionality of the original 22-item SERVQUAL, when adapted to IS, has been problematic. Due to the mixed results of Kettinger and Lee (1994) and of Pitt et al. (1995), it is important that further refinement of the IS SERVQUAL instrument be undertaken before it is confidently recommended as a prescriptive tool. To improve its practical value to IS managers, better definitions of the dimensions of IS service quality must be established. This can be accomplished by two approaches: (1) by careful reexamination of content validity, such as the development of underlying theory and the operationalization of items, and (2) by a parsimonious refinement using the specification search techniques of confirmatory factor analysis on the original 22 items.

Concerning the first approach discussed above, we would like to articulate our stance. While we concur with Pitt et al. (1997) that the domain of service quality seems to be adequately captured by the Parasuraman et al. (1988, 1991) original measures, we also agree with the suggestion by Van Dyke et al. that we should be mindful to the possibility that other
theoretically based dimensions of IS service quality may need to be included within SERVQUAL to accommodate different industries, cultures, and service contexts. However, while such additional theoretical and empirical work takes place, a pragmatic approach calls for continued improvement of SERVQUAL by refining our adapted measures to enhance item reliability and construct validity. Over the past several years in the marketing field, work has centered around this kind of improvement. It should also be noted that, even given the considerable debate in marketing concerning SERVQUAL, there have been few compelling challenges to its content validity.

To further clarify our position on content validity, we notice that Van Dyke et al. (1997) gave the false impression that the findings of Kettinger and Lee challenge the content validity of SERVQUAL. They claim that Kettinger and Lee (1994) found that neither User Information Satisfaction (UIS) nor SERVQUAL capture all of the factors which contribute to "Perceived Service Quality" in the IS domain. In fact, the aim of the 1994 study was to investigate whether a combined IS-adapted SERVQUAL and UIS measure better predicted "Summary UIS." That study did not examine the extent to which the combined SERVQUAL and UIS measure predicted "Overall Service Quality," which would be required to substantiate the claim by Van Dyke et al.

Establishing construct validity of the IS-adapted SERVQUAL

Discriminant validity is evident if items underlying a dimension load as discrete anticipated factors. This indicates the dimensions are then measuring different concepts. In the ideal case, exact reproduction of the five-factor model using the 22-item SERVQUAL instrument should indicate discriminant validity of the five dimensions. When tested, however, neither the Kettinger and Lee (1994) nor the Pitt et al. (1995) study established discriminant validity for the five Parasuraman et al. SERVQUAL dimensions.

In their study, Pitt et al. (1995) independently analyzed SERVQUAL data from three different sample sites using principal components and maximum likelihood methods with orthogonal rotation deriving three, five and seven factor solutions, respectively. In the seven-factor solution, the tangibles and empathy dimensions both split into two additional factors with reliability, responsiveness, and assurance dimensions loading close to the original SERVQUAL dimensions. The factor solutions for the other two samples did not load cleanly, nor closely resemble the original five SERVQUAL dimensions. Given their findings, Pitt et al. (1995) reported that "SERVQUAL does not clearly delineate among the dimensions of service quality" (p. 181). They warn users of the 22-item IS-adapted SERVQUAL instrument to be aware of the coalignment of the dimensions of responsiveness, assurance, and empathy due to the semantic similarity of these concepts and indicate that the reliability of the tangibles dimension is low.\(^2\)

In attempting to reconcile the SERVQUAL dimensionality problem, researchers, including ourselves, contend that much of this problem may be addressed by respecifying the items and deriving a refined instrument with clearer dimensionality and then cross-validating the derived model with an independent sample. In this vein, our approach has been to modify and deliver to practice a service quality measurement instrument better suited to the IS setting. Using the study data from our 1994 study, we found that a second order confirmatory factor analysis of the 22-item IS-adapted SERVQUAL resulted in poor discriminant validity. To derive a more practical instrument with more parsimonious distinctions between the five original dimensions, these 22 items were further refined by detecting and correct-

\(^2\)One problem that could have confounded the 1995 statistical analysis of Pitt et al. and resulted in poor dimensionality may be their use of an orthogonal factor rotation method for factors that have in past research been known to be highly correlated. The use of orthogonal rotation in light of known, highly correlated dimensions can result in unclear factor loadings (Stewart 1981). Previously, SERVQUAL studies have recognized high dimensional association and used oblique rotations or LISREL confirmatory factor analysis techniques.
ing specification errors (Andersen and Gerbing 1988; MacCallum 1986). While this type of refinement using a specification search has been criticized by some as exploratory model fitting, cross-validation of a derived model with an independent sample (Doll et al. 1995; Segars and Grover 1993) with independent observations (Chin and Todd 1995) substantially alleviates this concern.

Similar to several studies (e.g., Pitt et al. 1995; Parasuraman et al. 1991), the tangibles items proved to be problematic and was dropped from the model. However, as indicated in Table 1, the resulting instrument, which has fewer items but clearer distinction between dimensions, consists of 13 items comprising four of the five Parasuraman et al. dimensions (hereafter called the "SERVQUAL short form"). Cronbach alpha values for the final four SERVQUAL dimensions averaged .875, with the reliability dimension at .875, responsiveness at .883, assurance at .818, and empathy at .895. Significant y coefficients confirmed the convergent validity of the four SERVQUAL dimensions. To assess the discriminant validity of the IS-adapted SERVQUAL (short form) instrument, a formal test of discriminant validity was performed (Bagozzi and Phillips 1982). The chi-square difference between the constrained model (all Ys were constrained to 1.0) and the final model was 123.33 (p < .001; d.f. = 4), indicating discriminant validity. Further confirmation testing indicated that none of the confidence intervals among any two correlation coefficients for the four dimensions equaled or exceeded 1.00 (Bentler 1992; Shipper 1995). A total of 13 items loaded into four dimensions of SERVQUAL, demonstrating strong support for construct validity of the measure.

In the past, most SERVQUAL studies, including Kettinger and Lee (1994) and Pitt et al. (1995), have not statistically compared factor structures across different organizational samples. As recommended, a preferred method of cross-validating an instrument's dimensionality is to compare one factor structure against factor structures in multiple samples (Chin and Todd 1995). To achieve this, multiple group analysis with LISREL (MGAL), a statistical technique frequently used in marketing, was employed.

Our refined measurement model (SERVQUAL Sample 1) was cross-validated with a separate sample from a different organization. The SERVQUAL instrument was distributed to 80 IS users in another business school (SERVQUAL Sample 2), resulting in 48 usable responses. In applying the refined SERVQUAL Sample 1 factor structure to the SERVQUAL Sample 2, the two groups were stacked for simultaneous estimation. Using this method, the results of the comparisons between the SERVQUAL Sample 1 and the SERVQUAL Sample 2 indicated that the refined IS-adapted SERVQUAL 13-item (short form) instrument measured the same underlying concept of IS service quality within both samples.

While these findings provide an important direction in adapting SERVQUAL into the IS setting, additional research is needed before we can confidently accept the IS-adapted SERVQUAL short form. The proposed factor structure and dimensions need to be cross-validated within additional IS settings using larger sample sizes before confident claims of external validity can be made. For example, future research might use MGAL to cross validate these results with the three separate samples of Pitt et al. (1995).

While further testing should be undertaken, these findings have several important practical implications. First, this research indicates that practitioners can, with increased confidence, use a refined IS-adapted SERVQUAL as a means to conveniently measure specific service quality dimensions. Second, while Pitt et al. (1995) reported problems in establishing the discriminant validity of the responsiveness,
Table 1. Derived 13-Item IS-Adapted SERVQUAL

<table>
<thead>
<tr>
<th>Expected IS Service Quality</th>
<th>Perceived IS Service Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reliability:</strong></td>
<td><strong>Reliability:</strong></td>
</tr>
<tr>
<td>E1. When excellent information services promise to do something by a certain time, they will do so.</td>
<td>P1. When our information services promises to do something by a certain time, it does so.</td>
</tr>
<tr>
<td>E2. Excellent information services will perform the service right the first time.</td>
<td>P2. Our information services performs the service right the first time.</td>
</tr>
<tr>
<td>E3. Excellent information services will provide their services at the time they promise to do so.</td>
<td>P3. Our Information services provides its services at the time it promises to do so.</td>
</tr>
<tr>
<td><strong>Responsiveness:</strong></td>
<td><strong>Responsiveness:</strong></td>
</tr>
<tr>
<td>E4. Employees in excellent information services will give prompt service to users.</td>
<td>P4. Employees in our information services give you prompt service.</td>
</tr>
<tr>
<td>E5. Employees in excellent information services will always be willing to help users.</td>
<td>P5. Employees in our information services are always willing to help you.</td>
</tr>
<tr>
<td>E6. Employees in excellent information services will never be too busy to respond to users' requests.</td>
<td>P6. Employees in our information services are never too busy to respond to your requests.</td>
</tr>
<tr>
<td><strong>Assurance:</strong></td>
<td><strong>Assurance:</strong></td>
</tr>
<tr>
<td>E7. The behavior of employees in excellent information services will instill confidence in users.</td>
<td>P7. The behavior of employees in our information services instills confidence in you.</td>
</tr>
<tr>
<td>E8. Employees in excellent information services will be consistently courteous with users.</td>
<td>P8. Employees in our information services are consistently courteous with you.</td>
</tr>
<tr>
<td>E9. Employees in excellent information services will have the knowledge to answer users' questions.</td>
<td>P9. Employees in our information services have the knowledge to answer your questions.</td>
</tr>
<tr>
<td><strong>Empathy:</strong></td>
<td><strong>Empathy:</strong></td>
</tr>
<tr>
<td>E10. Excellent information services will give users individual attention.</td>
<td>P10. Our information services gives you individual attention.</td>
</tr>
<tr>
<td>E11. Excellent information services will have employees who give users personal attention.</td>
<td>P11. Our information services has employees who give you personal attention.</td>
</tr>
<tr>
<td>E12. Excellent information services will have the user's best interests at heart.</td>
<td>P12. Our information services has your best interests at heart.</td>
</tr>
<tr>
<td>E13. The employees of excellent information services will understand the specific needs of their users.</td>
<td>P13. Employees of our information services understand your specific needs.</td>
</tr>
</tbody>
</table>
assurance, and empathy dimensions, we found them to be construct valid and, as such, they should provide improvement in IS-adapted SERVQUAL’s diagnostic capabilities.

IS-Adapted SERVQUAL and Gap Scoring: A Matter of Managerial Choice?

Van Dyke et al. state major difficulties with SERVQUAL (Perception-Expectation) gap scores stemming from the operationalization of a service quality construct that was theoretically grounded in a discrepancy model. Based on work by marketing researchers that advocate direct perceptual measures instead of gap scores (e.g., Brown et al. 1993; Cronin and Taylor 1992, 1994), they contend that one’s perception of service quality already entails an expected service. By separately measuring the expected and perceived level of service quality and subtracting these scores, they argue, SERVQUAL is too simplistic to measure this complex cognitive evaluation process. Rather, they posit that a direct measure of perceptions, such as SERVPERF (Cronin and Taylor 1992, 1994), is a preferable means to capture the discrepancy between expected and perceived service quality.

Second, Van Dyke et al. argue that SERVQUAL’s expectation construct is ambiguous. Based on the work of Teas (1993) and Boulding et al. (1993), they state that the expectation measures suffer from multiple interpretations depending on whether a customer bases his/her assessment on a prediction of what will occur in the next IS service encounter or on what ideally should occur. To partially avoid the problem of an ambiguous expectation measure, researchers (Teas 1993; Peter et al. 1993) argue for a direct single-item comparative measure of the perception-expectation gap.

Although both Kettinger and Lee and Pitt et al. (1995) structured their instruments on the “ideal” expectation format using “excellent service” as an ideal anchor (based on Parasuraman et al. 1991 and Zeithaml et al. 1990, respectively), we concede that, after a careful examination of these items, it is possible that IS customers may rely on a different interpretation of expectations as a standard in making their expectation judgments. However, one might argue that the same confusion is possible in identifying one’s perceived level of service quality, which SERVQUAL critics argue is the implicit discrepancy between expectation and perception. That is, in direct perceptual measurement (SERVPERF), one may never know which interpretation of the “expectation” measure is assumed in the response.

Parasuraman et al. (1993; 1994a) provide careful rebuttals to the “gap score critics” and their arguments have been well characterized by Pitt et al. (1997). In addition to the conceptual justifications presented by Pitt et al. (1997), we believe that a vital way to judge the psychometric superiority of the “perception only” measure is to provide empirical evidence through direct comparison of SERVQUAL and SERVPERF in an IS setting.

In this regard, Lee and Kettinger (1996) conducted an empirical comparison between the original 22-item IS-adapted SERVQUAL instrument (Parasuraman et al. 1991) and a 22-item IS-adapted SERVPERF in terms of the psychometric superiority (refer to Table 2). Interestingly, this analysis showed very similar results to the later work of Parasuraman et al. (1994a). That is, it was determined that the SERVPERF instrument offered minor improvements in terms of reliability. However, the coefficient alpha values in both cases were consistently high across all dimensions, except for the tangibles dimension. When predictive validity was tested by a stepwise regression using the two service quality measures as independent variables, SERVPERF produced a modestly higher Adjusted R² value than SERVQUAL. The higher predictive validity of SERVPERF is consistent with previous marketing studies (e.g., Cronin and Taylor 1992; Parasuraman et al. 1994a) and supportive of the recent findings of Pitt et al. (1997). Next the 22-item IS-adapted SERVQUAL and the 22-item IS-adapted SERVPERF instrument were examined in terms of construct validity.
Research Note

It was found that SERVPERF did not show any evidence of better construct validity over SERVQUAL (refer to Table 2). In fact, neither the 22-item IS-adapted SERVQUAL nor the SERVPERF model was a good fit.

In sum, we found, as Van Dyke et al. (1997) proposed, that the predictive power of the IS-adapted SERVPERF instrument is superior to the IS-adapted SERVQUAL. Other comparative criteria, such as reliability and convergent and discriminant validity, show that the IS-adapted SERVPERF provides either weak or no substantial improvement over SERVQUAL. This finding is consistent with empirical evidence from marketing, where they have not conclusively established that SERVPERF is superior in terms of convergent or discriminant validity (Parasuraman et al. 1993, 1994).

Assuming that the findings from Lee and Kettinger and Pitt et al. (1997) are correct, IS managers might rightfully question: Why should we use SERVQUAL, which requires double the number of items, if SERVQUAL has no better psychometric properties than SERVPERF? In fact, Parasuraman et al. (1994a) and Pitt et al. (1997) address this question by suggesting that the richer information contained in SERVQUAL’s disconfirmation-based measurements provides managers with diagnostic power that typically outweighs the statistical and convenience benefits derived from the use of SERVPERF. Parasuraman et al. (1994a, p. 116) rhetorically ask, “Are managers who use service quality measurements more interested in accurately identifying service shortfalls or explaining variance in an overall measure of perceived service quality?” Ultimately, managers must decide based on their own unique needs whether SERVQUAL’s superior diagnostic value is preferable to SERVPERF’s convenience.

Moving beyond the SERVQUAL vs. SERVPERF debate

Since the introduction of SERVQUAL to IS, important work has transpired in marketing to better understand and improve the operationalization of expectation constructs. Significant contributions have come from the research efforts of Zeithaml et al. (1993) and Parasuraman et al. (1994b). In these studies, the authors acknowledge that their original “definition of the expectation construct was broad and did not stipulate norms of expectations used by customers in assessing quality” (Zeithaml et al. 1993, p. 3).

To better understand the differing standards of expectations as well as clarify the distinction between customer satisfaction and service quality assessment, Zeithaml et al. (1993)

Table 2. Comparison of IS-Adapted SERVQUAL and SERVPERF

<table>
<thead>
<tr>
<th>SERVQUAL</th>
<th>Tangibles: 0.668</th>
<th>Reliability: 0.880</th>
<th>Responsiveness: 0.861</th>
<th>Assurance: 0.830</th>
<th>Empathy: 0.870</th>
<th>Adjusted R2: Estimate: 0.455</th>
<th>p-value: 0.001</th>
</tr>
</thead>
<tbody>
<tr>
<td>SERVPERF</td>
<td>Tangibles: 0.816</td>
<td>Reliability: 0.878</td>
<td>Responsiveness: 0.867</td>
<td>Assurance: 0.851</td>
<td>Empathy: 0.861</td>
<td>Adjusted R2: Estimate: 0.579</td>
<td>p-value: 0.001</td>
</tr>
</tbody>
</table>

*X as a base model for convergent and discriminant validity testing, two CFA models were tested and no adequate fits were found respectively at the suggested cut-off values by Segars and Grover (1993).
combined past conceptualizations of service quality expectations with findings from a multi-sector study to develop an integrative model of customer service. As outlined in Figure 1, they compare customer evaluations of perceived service quality with customer evaluations of perceived customer satisfaction. Like Oliver's (1993) model of service quality-satisfaction, the Zeithaml et al. model distinguishes perceived service quality and perceived customer satisfaction as discrete constructs. Consistent with the marketing literature's customer satisfaction/dissatisfaction paradigm, the Zeithaml et al. model defines customer satisfaction as the difference between predicted service (what a customer believes will occur) and the perceived service (what a customer believes actually did occur).

However, in an effort to address criticism that SERVQUAL's expectation measure are ambiguous, Zeithaml et al. present a newer model (see Figure 1) that attempts to clarify the relationships among different standards of expectations and perceived service quality. This model posits that service expectations exist at two levels which customers can use as standards in assessing service quality: (1) desired service: the level of service representing a blend of what customers believe "can be" and "should be" provided, and (2) adequate service: the minimum level of service customers are willing to accept. Separating these two levels is a "zone of tolerance" that represents the range of service performance a customer would consider satisfactory. In other words, customer service expectations are characterized by a range of levels (between desired service and adequate service), rather than a single expectation point. Figure 2 depicts the zones of tolerances and perceived service quality levels of an imaginary company "A." By providing precise information about the perceived service levels across dimensions relative to adequate and desired service levels, insight is gained concerning the amount of emphasis that should be placed on improving poorly evaluated dimensions.

4A recent empirical study by Spreng and Mackoy (1996) supports this type of model and verifies that service quality and customer satisfaction are discrete constructs. Clearly, the concept of tolerance zones has the potential to overcome many of the complaints waged by Van Dyke et al. concerning the ambiguity of the IS-adapted SERVQUAL gap measure. In addition, this "desired" expectation, "adequate" expectation, and "perceived" service format, hereafter referred to as SERVQUAL+, has strong practitioner appeal. IS service providers may formulate service improvement plans based on this zone of tolerance concept. Depending on the relative position of the perceived service pointer (within or outside of the zone) for each of the different dimensions, short-term and long-term quality improvement resource allocation plans can be prescribed. In the short term, any dimension with a perceived pointer outside of the zone (such as responsiveness in Figure 2.1) would be a service dimension requiring utmost attention. When all of the pointers are within the zones, the relative positioning of any pointer and the width of the zone itself are criteria for deriving long-term IS service improvement plans.

IS providers serve multiple departments and customers with different zones of tolerance. Firm-wide evaluation may help prioritization of IS resources among these differing departments and customer types. For example, Figure 2.2 illustrates differences between the marketing and accounting departments in terms of zone widths, which suggests that marketing should receive closer attention because their zone of tolerance is narrower than accounting's, even though both possess similar positioning of perceived service pointers.

Depending on the type of IS service or customer (e.g., such as to an internal versus an external customer), the optimal positioning of a pointer within the zone may vary. For an internal IS customer, the lower end of the zone (close to adequate service) may be an optimal positioning of perceived service level, but this may not be acceptable for important external IS customers. Since a customer's desired service level is likely to be more stable than their perceived level of adequate service (Parasuraman et al. 1993), IS service providers may want to lower adequate service level expectations to maintain the widest pos-
Figure 1. Comparison Between Customer Evaluation of Perceived Quality and Satisfaction (Adapted from Zeithaml et al. 1993)
Figure 2. Illustration of “Zone of Tolerance” (Adapted from Parasuraman et al. 1994b)
sible tolerance zone. For example, clearly communicating the hours of operation of a user help desk as being 8 a.m. to 8 p.m. Monday through Saturday may help lower customers’ expectations for 24-hour/seven-days a week service. On the other hand, particularly for strategic customers (such as external customers of a home computer banking service), IS providers may consider raising the adequate expectation level to create a narrower tolerance zone in the hopes of keeping entrance barriers high for competitors. In such a case, the service agreement with the strategic customer might demand immediate, 24-hour/seven-days a week user support, thus raising the stakes for other IS service providers to offer the same.

In addition to managing tolerance zones, understanding the determinants of adequate service levels is vital in helping IS practitioners devise ways to manage expectations. Determinants of adequate service expectations might include a user’s past experience with the IS service, word-of-mouth, IS promises, and external vendors’ claims. Educating, communicating, and informing IS users about IS services and/or products, and realistic capabilities, are a necessary part of any IS marketing strategy. Without aggressive expectation management by IS service providers, users may tend to settle closer to their desired expectation points with cries for “more, faster, and better,” resulting in a narrower tolerance zone and a greater chance of perceived service quality failure.

Incorporating the model and concepts described above, Parasuraman et al. (1994b) proposed and tested three alternative service quality measures: (1) SERVQUAL+, which has a three-column format generating separate ratings of desired, adequate, and perceived service with three identical, side-by-side scales; (2) a two-column format, which generates direct ratings of the gap between desired and perceived service and the gap between perceived and adequate services; and (3) a one-column format, which measure direct ratings of the gap between desired service and perceived service level. While the basic content and structure of the original SERVQUAL instrument remains intact, a few major revisions and refinements were made. The proposed three-column, two-column and one-column approaches, adapted to IS, are shown in Figure 3.

The findings of Parasuraman et al., based on psychometric testings on these measurements, show very similar results to previous SERVQUAL and SERVPERF studies. The one-column format, which is an improved version of SERVPERF (with explicitly stated expectation standards), still demonstrates better predictive validity, while the other formats perform as well as the one-column format on all other psychometric criteria. In comparing the diagnostic value of the three formats, Parasuraman et al. (1994b) conclude that the three-column format is superior as a diagnostic tool because it provides detailed data to pinpoint service deficiencies and enough information to guide appropriate improvement efforts. These newer findings (Parasuraman et al. 1994b; Zeithaml et al. 1993) illustrate the potential for continuous improvement of the IS-adapted service quality measure and for utilizing the expectation construct as a diagnostic tool.

The most fruitful outcome of the long debate over SERVQUAL and SERVPERF in the marketing area is not a preponderance of psychometric evidence or theoretical justification indicating one approach is superior to another. Instead, there has been a gradual improvement of measures and their enhanced practicality in real world settings. Based on the research evidence presented here, Figure 4 highlights a comparison of SERVQUAL versus SERVPERF based on selected evaluation criteria. In this diagram, the overall pattern of findings from IS and marketing are generally summarized for managers to show tradeoffs in choosing a service quality scale. As can be observed, the varied placements of arrows indicates the relative advantage between SERVQUAL and SERVPERF on each criteria.

Further testing of these scales, and the adaptation of the promising SERVQUAL+ instrument to IS, will, most probably, result in improved instrumentation that possesses bet-
### Three-Column Format (SERVQUAL+)

<table>
<thead>
<tr>
<th>When it comes to...</th>
<th>My Minimum Service Level is:</th>
<th>My Desired Service Level is:</th>
<th>My Perception of ___'s Service Performance is:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Prompt service to IS users</td>
<td>Low 1 2 3 4 5 6 7 8 9</td>
<td>Low 1 2 3 4 5 6 7 8 9</td>
<td>Low 1 2 3 4 5 6 7 8 9</td>
</tr>
</tbody>
</table>

### Two-Column Format

<table>
<thead>
<tr>
<th>When it comes to...</th>
<th>Compared to My Minimum Service Level, ___'s Service Performance is:</th>
<th>Compared to My Desired Service Level, ___'s Service Performance is:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Prompt service to IS users</td>
<td>Lower 1 2 3 4 5 6 7 8 9 The Same 4 5 6 The Higher 7 8 9 No Opinion N</td>
<td>Lower 1 2 3 4 5 6 7 8 9 The Same 4 5 6 The Higher 7 8 9 No Opinion N</td>
</tr>
</tbody>
</table>

### One-Column Format

<table>
<thead>
<tr>
<th>When it comes to...</th>
<th>Lower Than My Desired Service level</th>
<th>The Same As My Desired Service level</th>
<th>Higher Than My Desired Service level</th>
<th>No Opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Prompt service to IS users</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td>4 5 6 7 8 9</td>
<td>8 9</td>
<td>N</td>
</tr>
</tbody>
</table>

Figure 3. Alternative Service Quality Measurement Formats (Adapted from Parasuraman et al. 1994b)
Conclusions and Future Directions

Van Dyke et al. have made an important contribution by bringing an ongoing debate from marketing to the IS community. We, however, stress the need for a practical point of view that is careful not to lose sight of the big picture of the IS service quality research paradigm by expediently retrofitting arguments from one group of marketing researchers into the IS community. Specifically, we take the position that measures should not be discarded until such time as their underlying theory and practicality for IS have been conceptually and empirically discredited. Until that time, IS researchers should concentrate on improving IS service quality measures.

Such “process improvement” is a lofty, scholarly goal, but one ultimately aimed at delivering useful tools to business practice. By offering the SERVQUAL short form, we counter the position of Van Dyke et al. that the IS-adapted SERVQUAL is not construct valid. However, we are not content to stop there. As did Parasuraman, Zeithaml and Berry in marketing, we intend to continue to work to improve the IS-adapted service quality measure. As illustrated in Figures 2 and 3, new formats, such as the SERVQUAL+ three column format and its “zone of tolerance” diagnostics, offer great potential for IS adaptation to practice. In addition, the improved SERVPERF (one-column format) has potential as an improved short form.

As depicted in Figure 4, there are discernible differences and relative advantages in using one scale or another. We believe that managers must select the service quality measurement that best meets their needs. With further empirical testing, managers may find that (1) SERVPERF is best for ease of application, situations requiring high predictive power, and possibility for a total service quality score; (2) SERVQUAL, including the short form, may be most applicable when greater dimensional diagnostics are required; (3) although longer in length, SERVQUAL+ may prove to be an improved form over SERVQUAL and SERVPERF. With its measures of zones of tolerance, SERVQUAL+ appears to have the greatest potential as an IS service quality diagnostics tool. Future IS research that empirically examines the rigor and relevancy of these measures will ultimately determine whether they can reach their potential.

In addition to empirical tests to refine the measures discussed above, future research should:

- identify the determinants of IS expectation levels and the extent to which they are unique;

- determine how specific tactics of expectation management affect IS user’s expectation levels;

- understand what is an optimal level of IS expectation management and whether overzealous management of expectations has negative effects on individuals and firm performance; and finally,

- examine whether standardized zones of tolerance can be used as industry-based benchmarks by which IS service providers can be measured.

We agree with the final comments of both Van Dyke et al. and Pitt et al. (1997): the IS service quality debate is just beginning. Such debate is fairly novel in IS. As an interdisciplinary field, IS research has too often adopted theories and measurements from reference disciplines only to quickly discard them for a more appealing approach without much discussion. We hope that the debate in this issue of MIS Quarterly
Figure 4. Manager's Choice: Relative Advantage of SERVQUAL vs. SERVPERF
signals a turning point in our research tradition, whereby different research teams attack a problem, engage in healthy scholarly interaction, and, in the process, help build a cumulative body of knowledge. Ultimately, the results of such focused research competition may result in the IS field's ability to contribute back to the referring discipline.

Acknowledgements

We would like to thank Bob Zmud and an anonymous associate editor for valuable comments improving this manuscript. Special thanks also go to Bonnie Glassberg for a careful read. This research has been partially supported by the Center of Information Management and Technology Research at the University of South Carolina and the Research Fund of the Perdue School of Business at Salisbury State University. Both authors contributed equally to this work.

References


### About the Authors

**William J. Kettinger** is director of the Center of Information Management and Technology Research at the University of South Carolina. He teaches in the Master's of International Business Studies program and in executive development programs both domestically and aboard. In addition to being on the faculty of USC for many years, Dr. Kettinger has over 18 years professional consulting experience in both the public and private sectors with such organizations as IBM, NCR, Bose, Philips N.V., and the South Carolina state government. Between 1992 and 1994, he assisted AT&T's Software Division in developing a business process management strategy. More recently he has been advising the *Business Gateway* of the Enterprise Development Institute in building an Internet-based electronic bid matching and EDI capability for small businesses. His current consulting and research focuses on electronic commerce.
strategy, business process change, and IS service quality. He has published in such journals as *MIS Quarterly*, *Decision Sciences*, *Journal of Management Information Systems*, *Communications of the ACM*, *Public Administrative Review*, *Data Base*, and *Information & Management*. He has served as a past special editor for *Journal of Management Information Systems* and co-edited the book *Business Process Change: Concepts, Methods and Technologies*. He was a 1995 recipient of the Society for Information Management's Best Paper Award. He received his Ph.D. and an M.S. from the University of South Carolina and an M.P.A. from the University of Massachusetts at Amherst.

Choong C. Lee is an assistant professor of MIS at the Franklin P. Perdue School of Business at Salisbury State University. He has a Ph.D. in MIS from the University of South Carolina, an MBA from the University of Rhode Island, and a B.A. degree from Yonsei University, Seoul, Korea. Dr. Lee has published in such journals as *Decision Sciences*, *Journal of Management Information Systems*, *Data Base*, *Information and Management*, and *Information Resources Management Journal*. His research interests include strategic telecommunication planning, IS service quality, IT impact on strategy and organization structure, IS curriculum design, and IS performance measurement.