Measuring Information Systems Service Quality: Concerns for a Complete Canvas

By: Leyland F. Pitt
Cardiff Business School, University of Wales College of Cardiff, Aberconway Building, Colum Drive
Cardiff CF1 3EU
UNITED KINGDOM
pittl@cf.ac.uk

Richard T. Watson
Department of Management
University of Georgia
Athens, GA 30602-6256
U.S.A.
rwatson@uga.cc.uga.edu

C. Bruce Kavan
Department of Management, Marketing, and Logistics
University of North Florida
4567 St. John's Bluff Road
Jacksonville, FL 32216
U.S.A.
bkavan@unf.edu

Abstract

This paper responds to the research note in this issue by Van Dyke et al. concerning the use of SERVQUAL, an instrument to measure service quality, and its use in the IS domain. This paper attempts to balance some of the arguments they raise from the marketing literature on the topic with the well-documented counterarguments of SERVQUAL's developers, as well as our own research evidence and observations in an IS-specific environment. Specifically, evidence is provided to show that the service quality perceptions-expectations subtraction in SERVQUAL is far more rigorously grounded than Van Dyke et al. suggest; that the expectations construct, while potentially ambiguous, is generally a vector in the case of an IS department; and that the dimensions of service quality seem to be as applicable to the IS department as to any other organizational setting. Then, the paper demonstrates that the problems of reliability of difference score calculations in SERVQUAL are not nearly as serious as Van Dyke et al. suggest; that while perceptions-only measurement of service quality might have marginally better predictive and convergent validity, this comes at considerable expense to managerial diagnostics; and reiterate some of the problems of dimensional instability found in our previous research, highlighted by Van Dyke et al. and discussed in many other studies of SERVQUAL across a range of settings. Finally, four areas for further research in this area are identified.

Keywords: Measurement, reliability, validity, service quality, marketing of IS, IS research agenda

ISRL Categories: AI04, AI0402, AI0403, EI0206.03, EL01, IB04

Introduction

A research note in this issue (Van Dyke et al. 1997) comments on the use of SERVQUAL, an instrument developed by marketing academics (Parasuraman et al. 1985, 1988, 1991, 1993), to measure IS service quality (Kettinger and Lee 1994; Pitt et al. 1995). Specifically, Van Dyke et al. concentrate on the conceptual and empirical difficulties with
SERVQUAL and provide a sound summary of the marketing literature in this regard. From a conceptual perspective, after rightly questioning the use of subtraction as a simulation of a psychological process, they focus on the work of Teas (1993, 1994), on the ambiguous interpretation of the expectations side of the SERVQUAL equation, and that of Carman (1990), on SERVQUAL's general applicability. From an empirical perspective, they highlight the efforts of Peter et al. (1993), which also appears elsewhere (see Brown et al. (1993), on measuring reliability when difference scores are used; and various attempts (cf. Cronin and Taylor, 1992, 1994) to demonstrate that predictive validity might be better achieved by performance (or perceptions-only) based measures of service quality than by the use of the gaps obtained by subtracting expectations from perceptions.

The IS community needs to be aware of problems that might be experienced in using an instrument to measure so critical a construct as IS service quality, related as this construct may be to other issues of crucial concern, such as user satisfaction, effectiveness and productivity. However, the IS community is also worthy of a comprehensive picture, a fully-fledged (or at least nearly finished) painting, a completed canvas, as it were. The SERVQUAL debate has certainly not been one-sided. Parasuraman et al. have responded in full, with sound and solid arguments regarding the conceptual and empirical aspects of their conceptualization of service quality. The IS community deserves to hear these counterarguments.

We attempt to add further color to the canvas regarding the conceptual and empirical difficulties identified by Van Dyke et al. by summarizing some of the responses to these difficulties. Then the specific problems of using SERVQUAL within IS and the use of SERVQUAL in these settings are addressed. The paper concludes by identifying some areas for future research and debate.

Conceptual problems

Van Dyke et al. rightly point out that it is difficult to distinguish clearly between conceptual and empirical problems in this type of research. Following their arguments, this paper addresses their concerns in the same sequence.

Subtraction as simulation

Van Dyke et al. reiterate the concerns of marketing researchers (e.g. Cronin and Taylor 1992) that conceptualizing service quality as the gap between a customer's perceptions and expectations is problematic, over-simplistic, and lacking in theoretical or empirical evidence. This seems to negate a very strong and recurring theme in the Parasuraman et al. research. The gap formulation was not merely suggested or implied as a convenient simplification, it was derived and conceptualized from extensive focus group research (Parasuraman et al. 1985; Zeithaml et al. 1990); grounded in prior conceptual work cited by Pitt et al. (1995), namely Gronroos (1982) and Sasser et al. (1978); substantially supported by Bolton and Drew (1991a) work cited by Van Dyke et al.); and reaffirmed by the same authors (Bolton and Drew 1991b, p. 383) (in a work not cited by Van Dyke et al.) as follows:

Consistent with prior exploratory research concerning service quality, a key determinant of overall service quality is the gap between performance and expectations (i.e., disconfirmation).

Van Dyke et al. broadly criticize this subtraction operationalization of the disconfirmation paradigm as being over-simplistic, without really offering feasible alternatives. Merely suggesting that direct measurement of perceptions is better, or that some kind of "direct measurement of the discrepancy" is more desirable does not leave us better off. These approaches are well known in the marketing literature (cf. Cronin and Taylor 1992) and might result in marginally better predictive validity. The diagnostic insights they provide to the IS manager are questionable.
Is the expectations construct ambiguous?

Van Dyke et al. summarize the work of Teas (1993, 1994) on the use of expectations measures in service quality research. This paper certainly concedes that there do not appear to be readily available solutions to all the problems that seem to be associated with precisely defining what an expectation of service quality is. Teas (1993) points that the SERVQUAL conceptualization of expectations is as a vector—"in simple terms, more is always better."

Van Dyke et al. assert that this causes the problem of the expectation score becoming less useful in subsequent analysis as an increasing proportion of the variation in difference-based SERVQUAL is due only to changes in perceptions, and also point out that the conceptualization of the SERVQUAL expectations items as will statements (Pitt et al. 1995) is opposite from that intended by the SERVQUAL authors, according to Boulding et al. (1993). However, the following points are worth noting. First, Parasuraman et al. (1994) state that it is likely that customers consider most of the 22 items in the SERVQUAL instrument to be vector attributes (a point also acknowledged by Teas in footnote 16 of his article). Second, in their response to Brown et al. (1993), who also noted the variability problem in a similar fashion, Parasuraman et al. (1993) argue that this will only be really problematic when the SERVQUAL gaps are used as dependent variables in analysis. Third, the conceptualization of the SERVQUAL expectations items (Pitt et al. 1995) was entirely consistent with that of Zeithaml et al. (1990) and followed the wording of the questionnaire, with minor alterations to cater for the IS department, outlined in the appendix to their book. We anchored the respondents on excellence in the preamble and instructions, asking them to "think about the kind of IS department that would provide excellent service," and then proceeded, as Van Dyke et al. rightly point out, to use statements reading "they (the IS department) will. . . ." Kettinger and Lee (1994) used statements that all included the word excellent, and phrased along these lines: "excellent . . . will. . . ."

Van Dyke et al. also point to Teas' (1993) reasoning that an expectation can be interpreted as a classic ideal point rather than a vector. This point is conceded by Parasuraman et al. (1994) in their response to Teas. Although, as briefly discussed in the previous section, Parasuraman et al. believe that customers will see most of the 22 SERVQUAL items as vectors, we think that, in the IS environment, there may be a few expectation items in SERVQUAL that could be classic ideal points. Indeed this might in part account for the instability of the tangibles dimension, noted by a number of researchers (Kettinger and Lee 1994) and ourselves. IS customers might be dissatisfied by IS employees who are too neat in appearance (SERVQUAL item 3), while still delivering inadequate service in terms of dimensions such as reliability, assurance and responsiveness, or when the appearance of the physical facilities of the IS department is more than in keeping with the kind of services provided (SERVQUAL item 4), particularly when their own facilities do not match these.

Finally, Van Dyke et al. seem to have misunderstood some of the implications of the work of Boulding et al. (1993)—e.g., "These findings fail to support the (P-E) gap model." It is important to note that Boulding et al. did not test the gap formulation against their own formulation (i.e., E affecting P which, in turn, affects SQ). As such, just because their findings support links from E to P to SQ does not necessarily imply the P-E gap formulation is problematic.

Applicability of SERVQUAL across industries and settings

Repeating the arguments of other researchers (Carman 1990; Brown et al. 1993; Dabholkar et al. 1996), Van Dyke et al. question the universal applicability of SERVQUAL as a service

---

2Contrary to the assertion by Van Dyke et al. that we give a blanket go-ahead to the use of SERVQUAL in IS, we state that "Potential users of SERVQUAL should be cautious," among other reasons because the reliability of the tangibles construct is low (Pitt et al. 1995; p. 181).
quality measure, and specifically its use for IS departments. Since they have omitted the well-considered response of Parasuraman et al. (1993) to Brown et al., we begin by balancing the argument.

Parasuraman et al. (1991) assert that the SERVQUAL items represent core evaluation criteria that transcend specific companies and industries, providing a basic skeleton underlying service quality that can be supplemented with context-specific items when necessary (and they provide instructions for doing so). They go on to point out that items seen as different or missing by Brown et al. (convenient operating hours) are in fact contained in SERVQUAL, in item 19.

Van Dyke et al. raise a key question, however: Are the services provided by the IS department different enough to warrant a separate instrument for the measurement of its service quality, or at least the addition of items to the standard SERVQUAL? In principle (following Parasuraman et al. 1991), we have nothing against the latter should this be warranted, and certainly stated nothing to the contrary in Pitt et al. (1995). Moreover, we provide three reasons meriting the use of a generic measure.

First, as Van Dyke et al. point out, we cannot discern any unique features of IS that make the standard SERVQUAL dimensions inappropriate. It is unlikely that there will be IS departments that do not wish to strive to provide service that is right the first time, keep their promises (reliability), be prompt and willing to help (responsiveness), be courteous, competent and offer security (assurance), treat customers as individuals (empathy), and have equipment, facilities, printed materials, and people of appropriate appearance (tangibles).

Second, we cannot discern any unique features in the IS domain that have been excluded from SERVQUAL, although that does not mean we are not open to suggestion. Many of the issues that both managers and researchers raise, we feel, can with some consideration be reassorted under a SERVQUAL dimension. To use the Kettinger and Lee (1994) extractions from the Baroudi and Orlikowski (1988) UIS as examples (as pointed out by Van Dyke et al.):

1. Degree of training provided to users by IS staff might be an assurance item;

2. Level of communication between users and IS staff might be an empathy item;

3. Time required for new systems development and implementation might be a reliability item (indeed our SERVQUAL item 5 reads, “When IS promises to do something by a certain time, it does so”).

Obviously in the absence of data and appropriate analysis, this may be speculation. However, we do think there is considerable overlap between the UIS dimensions cited by Van Dyke et al. and summarized by them as training, communication and time, and the SERVQUAL items and dimensions. For example SERVQUAL uses words such as time, when as an indication of time prompt, hours, and always, which appear in a number of items.

Third, there is a good diagnostic reason for using a generic measure of service quality. While IS managers may want to know how they compare to other IS departments, nowadays many want to compare themselves to other excellent service providers, especially those serving external customers. As one senior executive in the IS department of a major airline that had used SERVQUAL told us recently,

I know we are one of the best IS departments around, because so many other IS departments in other businesses benchmark on us. That’s nice, but it’s not enough for us. We want to know how we compare to great service providers who do not enjoy the monopoly position that we do—our own airline’s passengers, for example, the best credit card companies, top hotels.

SERVQUAL provides an opportunity for this on a cross-industry and cross-functional basis, being a widely used, generic instrument.
Empirical Problems

Van Dyke et al. reiterate the three main points of debate that have occurred in the marketing literature: the reliability issues associated with the use of difference scores in measurement; whether a difference score results in lower predictive and convergent validities than would a performance based measure; the unstable dimensionality of SERVQUAL.

Reliability of difference scores

Van Dyke et al. refer to two well conceptualized articles (Brown et al. 1993; Peter et al. 1993) that highlight the problems associated with using a difference score to operationalize a construct. Simply, operationalization of a construct using a difference score means that the construct (for example, the service quality gap, G) is calculated as the difference between one more construct (such as the perceptions of service quality, P), and another (the customer's expectations of service quality, E), i.e., \( G = P - E \). Peter et al. discuss the issues that this raises in detail. This work merits the attention of researchers beyond the marketing community. Brown et al. focus particularly on the difference score conceptualization of SERVQUAL, and this piece was the subject of a well-considered response by Parasuraman et al. (1993).

Most researchers, including Parasuraman, Zeithaml and Berry as well as ourselves, have used the standard formula for the calculation of coefficient alpha (Cronbach 1951) to calculate the reliability coefficients for the SERVQUAL dimensions. As Van Dyke et al. show from Brown et al. (who in turn cite Johns 1981), this may be inappropriate. The real question is, of course, whether the effect of this is serious.

Brown et al. correctly state that the reliability of a scale operationalized as the difference between two measures will be low to the extent that (1) the correlation between the component measures is high (in this instance, perceptions and expectations) and (2) the reliabilities of the component measures (perceptions and expectations) is low. Parasuraman et al. (1993) argue strongly (and demonstrate with the data of Brown et al.) that these conditions (1) and (2) are not likely to be serious problems in the case of SERVQUAL however, where the construct G is operationalized as P-E.

If we pursue this argument in an IS context, there is really no good conceptual reason why expectations of IS service and perceptions thereof should be highly correlated. A user having higher expectations than another of the IS department's knowledge to do the job well does not by definition, therefore, also have higher perceptions of the IS department's knowledge to do the job well. The data of Brown et al. does not show high correlation between expectations and perceptions (.34), neither does their data according to Parasuraman et al. (1993), and our own correlations are .31, .21, and .21 for the three firms we studied. Furthermore, condition (2) above does not seem to generally hold in the case of SERVQUAL—Brown et al. report high reliabilities for the expectations (.94) and perceptions (.96) components of the measure, an observation confirmed by Parasuraman et al. (1993) from their work. Our calculation of the reliabilities for the components in Pitt et al. (1995) give relatively high reliabilities for expectations (.76, .93, and .97) and perceptions (.81, .96, and .97). We also recalculated the reliabilities using the formula referred to by Van Dyke et al., from Brown et al.

The reliability of the difference scores when calculated using the component variables are quite close to those originally calculated (see Table 1), with the largest difference being .05. The recalculation has four of 15 reliabilities below .70, with three of these for tangibles. In Pitt et al. (1995), we report three of 15 reliabilities below .70, and two of these are for tangibles.

Finally, Van Dyke et al. conveniently cite the findings of Parasuraman et al. (1994) for the computer manufacturer to point out that the difference-score formulation of SQ has lower reliability than the direct-measure formulation. However, they fail to acknowledge the discussion elsewhere in Parasuraman et al. (1994),
Table 1. Reliability of SERVQUAL by Dimension

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Number of Items</th>
<th>Financial Institution</th>
<th></th>
<th>Consulting Firm</th>
<th></th>
<th>Information Service</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tangibles</td>
<td>4</td>
<td>227</td>
<td>Gap</td>
<td>0.65</td>
<td>0.65</td>
<td>0.73</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Difference</td>
<td>0.58</td>
<td>0.65</td>
<td>0.73</td>
<td>0.69</td>
</tr>
<tr>
<td>Reliability</td>
<td>5</td>
<td>181</td>
<td>Gap</td>
<td>0.86</td>
<td>0.88</td>
<td>0.94</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Difference</td>
<td>0.86</td>
<td>0.88</td>
<td>0.94</td>
<td>0.93</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>4</td>
<td>267</td>
<td>Gap</td>
<td>0.82</td>
<td>0.83</td>
<td>0.96</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Difference</td>
<td>0.82</td>
<td>0.83</td>
<td>0.96</td>
<td>0.91</td>
</tr>
<tr>
<td>Assurance</td>
<td>4</td>
<td></td>
<td>0.83</td>
<td>0.84</td>
<td></td>
<td>0.91</td>
<td>0.87</td>
</tr>
<tr>
<td>Empathy</td>
<td>5</td>
<td></td>
<td>0.82</td>
<td>0.82</td>
<td></td>
<td>0.90</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Gap is the value for reliability calculated using gap scores.
Difference is the value for reliability calculated using component variables.

which includes empirical evidence, that suggests that the direct measures may suffer from much higher measurement error (perhaps due to respondent confusion) than the difference-score measure, and from an upward bias (i.e., the direct measures may be painting a rosier—and misleading—picture than warranted by reality).

Predictive and convergent validity of SERVQUAL

Van Dyke et al. point to the prolific and ongoing debate (cf. Babakus and Boller 1992; Brown et al. 1993; Cronin and Taylor 1992; Parasuraman et al. 1991, 1993, 1994) concerning the predictive and convergent validities of SERVQUAL. The debate has not, as they assert in their opening sentence, centered on the poor validity of SERVQUAL (for it is arguably not poor). Rather, it has concentrated on the fact that, in the research of which we are aware, performance based or perceptions-only measures of service quality have (1) higher correlations with other measures of the same construct (convergent validity) and (2) higher correlations with other conceptually related constructs. Following Parasuraman et al. (1988), in our IS SERVQUAL research we have used an overall evaluation of IS department service quality as another conceptualization of the same construct (i.e., (1)), and items such as whether the respondent had reported a service quality problem to the IS department and whether this had been satisfactorily resolved as conceptually related constructs (i.e., (2)).

Perceptions-only measures of service quality generally have both higher convergent and predictive validities than SERVQUAL difference scores. Not only has this been widely established in the literature to which Van Dyke et al. refer, it has also been conceded by Parasuraman et al. on a number of occasions (1991, 1993, 1994). While we did not report the regression between the perceptions-only items and an overall measure of service quality in Pitt et al. (1995), we do so here.

As can be seen from Table 2, the perceptions-only items have a higher R² gaps measure in two of the three cases. Nevertheless, we agree with Parasuraman et al. (1994), in their response to Cronin and Taylor (1992), that service quality measurements that incorporate customer expectations provide richer diagnostic information than those that merely use perceptions. We concede that the perceptions-only measures may possess stronger statistical properties than do gaps measures. However we believe that IS managers find the concept of a gap most useful, for it indicates a shortfall, something that needs to be closed. Just as importantly, we believe that IS managers want to know and should know what their customers expect. Frequently, in discussions of IS service quality study results, a manager might say something like,
Table 2. Regression of the Overall Service Quality Measure with Item Perception and Gap Measures

<table>
<thead>
<tr>
<th>Financial Institution</th>
<th>Consulting Firm</th>
<th>Information Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perception Gap</td>
<td>Perception Gap</td>
<td>Perception Gap</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.50</td>
<td>.44</td>
<td>.40</td>
</tr>
<tr>
<td>.40</td>
<td>.69</td>
<td>.64</td>
</tr>
<tr>
<td>.39</td>
<td></td>
<td>.39</td>
</tr>
</tbody>
</table>

"Customers expect too much." How would they know this, or be able to test the assertion, if expectations were not measured? Both academics (cf. Berry and Parasuraman 1991) and practitioners are interested in the management of customer expectations. How could these be managed if they were not first measured? There might be important differences between various groups of customers with regard to their expectations of service quality, which might have customer relationship implications. These would not be detected if expectations were not measured. For example, in an IS-like setting, Pitt et al. (1992) found significant differences between the expectations of software experts (who emphasized reliability) and software users (who emphasized empathy) as customers of a mainframe software supplier in a SERVQUAL study, which would have been overlooked had a perceptions-only framework been used. Finally, measuring both expectations and perceptions over time gives a better indication of where changes have occurred (Parasuraman et al. 1993). We have demonstrated this in an IS context, in the case of an IS department where customer expectations had risen over time, but the overall gap had decreased because of an increase in customer perceptions of performance (Pitt and Watson 1994). In summary, without wishing to sacrifice statistical rigor, we are nevertheless reluctant to “throw the baby out with the bath water”—managers benefit from the additional diagnostics which a gap measure and its components provide. While perceptions-only based measures of service quality may have superior statistical properties, the gaps measurement yields richer managerial insights that it would be a great pity to lose.

**Dimensional instability**

It has been a source of ongoing concern, beginning with Carman (1990), and also in Pitt et al. (1995) and Kettinger and Lee (1994), that the factor structure of SERVQUAL is unstable and that the factors or dimensions defined by Parasuraman et al. (1988) do not generally emerge as such in subsequent research. Parasuraman et al. (1991) discuss the issue comprehensively and offer a comparative discussion of their findings and those of other researchers. Parasuraman et al. (1994) again acknowledge these issues and suggest, in a similar fashion to our proposition regarding Q-sort (Pitt et al. (1995), that a customer based, sorting assessment of the dimensionality should offer intriguing insights. Van Dyke et al. give an excellent summary of this discussion, and we have little to add except for one point.

Van Dyke et al. claim that there is no underlying “theory” to justify the combining of “understanding” and “access” into empathy. This claim is debatable. Clearly, one can argue that being accessible to and by customers is important for understanding their needs; alternatively, being accessible to customers may also be a consequence of a company’s truly understanding its customers and “putting itself in their shoes”—the core of what empathy stands for. Thus, here and elsewhere in Van Dyke et al., their claims about SERVQUAL’s being “atheoretical” are quite puzzling. Related to this, and somewhat ironically, Van Dyke et al. seem to suggest in their discussion section that IS professionals would be better off measuring and managing “will” and “should” expectations. Several questions are worth raising in this regard: What exactly is the theoretical basis for these two types of expectations? (The formulation of these two types of expectations by Boulding et
al (1993) was preceded by little, if any, of the extensive exploratory, conceptual work that preceded the SERVQUAL formulation.) Is not the support for these two types of expectations primarily empirically based, rather than theory based? Most importantly, what exactly do "will" and "should" expectations mean in an IS context, and how can IS professionals differentially measure and influence the two types of expectations? Important issues are raised but remain unaddressed by Van Dyke et al.

Some Avenues for Future Research on Service Quality in IS

Rather than suffice with this response to the important issues raised by Van Dyke et al., we go further by attempting to translate some of the preceding discussion into an identification of some areas for future research. Van Dyke et al. have brought many of the concerns regarding SERVQUAL identified in the marketing literature to the attention of the IS community, and hopefully we have in some way moved to complete the picture, but it would be a shame at this stage not to take this dialogue out of a mere problem-response debate and into a more fruitful research realm. We introduce four avenues of consideration for future research, which we believe are worthy of the IS academic's attention.

Are there areas of IS where expectations are not vectors?

While Parasuraman et al. (1994) assert that the SERVQUAL expectation items, and indeed dimensions, all seem to be vectors and while in our response we have agreed that it is difficult to conceive of them as otherwise, Van Dyke et al. may have a point. This may be particularly true within the IS domain (particularly if it is extended to include related measures of user satisfaction such as UIS). It would therefore seem appropriate to investigate the possibility that some IS service quality concerns may be classic ideal-points, where the expectation peaks and then declines, so that doing too much of it may result in customer dissatisfaction (e.g., IS staff are too well-dressed and too friendly, communication is too on-going and too frequent, and the brochures, manuals and newsletters are too glossy and too expensive). If this were found to be the case, then it raises both measurement and management issues: How does one manage service quality in an IS department where doing too much for the customer will result in dissatisfaction?

How do IS managers interpret SERVQUAL data, and what do they find useful?

Most of the concerns regarding the measurement of service quality have focused on the customer side of the equation. This is understandable and obviously important, but it is at this stage probably not enough. Marketing researchers such as Cronin and Taylor (1992) and Brown et al. (1993), echoed by Van Dyke et al., assert that perceptions-only based measures of service quality demonstrate greater convergent and predictive validities, while Parasuraman et al. (1993, 1994) remonstrate that managers benefit more from the greater diagnostics of gap measurement. This does not resolve the issue of what managers really want. While we believe that managers find the diagnostic power of expectations measurement to be more useful, we readily admit that this has not been the central focus of our research, and that what evidence we do have comes from a relatively small sample of IS managers who commissioned IS service quality studies. It therefore seems appropriate to investigate what managers want with regard to IS service quality measurement in IS.

The IS department is an internal marketer

Based on the work of Carman (1990) and Brown et al. (1993), Van Dyke et al. question
the universal applicability of SERVQUAL and, in particular, its generic applicability in an IS setting. We have responded that the IS department is a service provider and therefore the measurement of service quality should not be that different from other service providers. However, there is a possibility that both sides of the debate might be overlooking the issue that could matter most: That it is not that IS itself is unique, but that internal service provider departments in general might be different from providers of services to external customers. Thus, IS departments, training departments, housing and relocation functions, and all other entities that provide services or products only to the organization’s employees (internal customers) may be sufficiently different from other service businesses to warrant particular attention when it comes to measuring service quality.

These are unique situations, for in many instances the IS department is seen as an inflexible monopoly by its customers, and it in turn experiences the problems of being on a “hiding to nothing”—when things go well, it is because the users are so capable, and when things go wrong, it is always the IS department’s fault. These issues are brilliantly highlighted in San Francisco Bay Consulting (Baker and Monsler 1994), which features the travails of an IS department dealing with its ungrateful and over-demanding internal customers. Internal marketing is an issue that is becoming an increasingly important one in the marketing literature (cf. Bowen and Schneider 1985; Comm 1989; Fisk et al. 1993; George 1990; Harrell and Fors 1992; Lee et al. 1991; Piercy and Morgan 1992). It is an avenue that the IS research community may wish to follow, particularly in view of the ongoing drive toward IS outsourcing. The central issue is whether service quality measurement in an internal market is different from that in external markets, not because IS is different from other functions or industries, but because it is an internal, rather than an external, service provider.

**IS service quality measurement and the disconfirmation paradigm**

Van Dyke et al. alert the IS community to concerns voiced regarding the conceptualization of service quality as a difference. We responded by referring to the counterarguments of Parasuraman, Zeithaml and Berry in this regard, which are based in the disconfirmation literature (Oliver 1980). The most insightful criticism of an overall applicability of disconfirmation, in our opinion, comes from work in consumer research (Deighton 1992).

As Deighton points out, the expectancy-disconfirmation analysis of customer satisfaction (Oliver 1980) describes very well how judgments are formed of the quality of contracted performance—that is, performances in which the role of an actor is not salient. Thus when people use a particular brand of diskette and decide that it performs well, they do not usually think of the manufacturer, or of who procured it for them (the IS department) or even of their own skill in using it. In Deighton’s words, they simply credit the performance to the object and judge the performance’s quality against an internalized standard for similar objects. “Expectations correspond to the imputed obligation to perform, and confirmation or disconfirmation corresponds to discharge of the obligation or the failure to do so.” Similar evaluations could be made of the availability of a network, on-line help facilities, and many of the routine services IS departments provide.

When the customer recognizes that meeting the obligation toward them involves a person, performance becomes an enacted event (Weick 1979). Deighton believes that the expectancy-disconfirmation model of customer satisfaction does not fit this case very well. This is because there is room for latitude in the way the customer allocates credit or blame for the performance between actor and object. The actor can be the provider, the customer, or both. This situation might be typical of much of the service an IS department produces. On the provider side, for example, satisfaction with the performance of an IT-related service might be the result of an appropriate system-related
procedure (object) or an excellent execution of the procedure (actor). The marketing problem facing the IS department here would be to influence how the credit is allocated and what devices of impression management to use to affect the allocation (Deighton 1992). On the customer side, to take another example (Deighton's own), if a customer says that a computer software program performs poorly, perhaps she, as actor, has played some part in the poor outcome. If a novice computer user hears an expert praise a software product, he may well wonder if it will perform as well for him. Here, too, the craft of performance marketing deals with how these attributions are influenced.

What Deighton alerts us to is not so much that industries or functions differ, but that the various services they deliver may differ in terms of the nature of their performance. The disconfirmation paradigm does not explain quality equally well across different kinds of performance, according to Deighton. If we accept this for the services provided by the IS department, then the research perspectives offered by Deighton are considerable. The work of Zeithaml et al. (1993) on expectations also has several concepts that may offer further support for some of the arguments advanced here. For instance, "self-perceived service role" and "situational factors" (two of the hypothesized determinants of customers' expectations) may be relevant for, and augment, our arguments about customer attributions in interpreting SQ. Another concept in Zeithaml et al. that seems similar to Deighton's "imputed obligation" notion is "implied service promises."

Conclusion

Van Dyke et al. have done well to remind the IS community of the importance of the IS department as a service provider, the ongoing debate regarding the measurement of IS service quality in general, and particularly of the concerns regarding the use of SERVQUAL to measure IS service quality. They have alerted the IS community in general to the criticisms of the SERVQUAL instrument in the marketing literature, and have done well at directing these criticisms toward the use of SERVQUAL in IS. We hope that in our response we have balanced the picture by making readers aware of the counterpoints to these criticisms where appropriate and reaffirmed the problems where these exist. Not wishing to let a good argument end there, we have concluded by suggesting areas for research in IS which have been implicitly raised by this debate. No good canvas is completed in a single attempt, it tends to be the result of much effort and hard work. Hopefully, both the Van Dyke et al. contribution and our response will see the canvas move a small step nearer to being a clear picture.

Acknowledgements

We gratefully acknowledge the very helpful comments of A. "Parsu" Parasuraman, James McLamore Chair in Marketing, University of Miami. All errors and omissions remain our own.

References

Bolton, R. N., and Drew, J. H. "A Multistage Model of Customers’ Assessments of


### About the Authors

Leyland F. Pitt is professor of marketing and strategy at Cardiff Business School, University of Wales. He also teaches on executive development programs at the London Business School. He has held executive positions in industry and has taught marketing and strategy at other universities in Europe (UK, France, Portugal, Norway, Malta and Greece), Southern Africa, Australia, South-East Asia, Canada, and the USA, where he regularly lectures on executive programs at the Graham School of General Studies at the University of Chicago and the Graduate School of Business at Columbia University. His work has been accepted for publication by such journals as *The Journal of Advertising Research*, *The Journal of Business Research*, *Business Horizons*, *The Journal of Direct Marketing*, *Industrial Marketing Management*, *The Journal of Small Business Management*, and *MIS Quarterly*. His particular areas of interest in research and teaching include the future of brands and alternatives to product and brand manager systems; customer search and the new media; and the information technology–marketing interface. Pitt has consulted for firms and government departments in many parts of the world and has also been the recipient of a number of awards for teaching excellence. He holds MCom in management from
Rhodes University and an MBA and Ph.D. in marketing from the University of Pretoria.

Richard Watson is an associate professor in the Department of Management at the University of Georgia. His Ph.D. in management information systems was awarded by the University of Minnesota. He has published in leading journals in MIS, auditing, marketing, business ethics, and communication, as well as authored books on data management and electronic commerce. He is on the Editorial Board of MIS Quarterly and three other journals. His current research, which has a strong international flavor, focuses primarily on electronic commerce and management of the MIS function.

C. Bruce Kavan is an associate professor of MIS and interim director of the Barnett Institute at the University of North Florida. Prior to completing his doctorate at the University of Georgia in 1991, he held several executive positions with Dun & Bradstreet including vice president of Information Services for their Receivable Management Services Division. Dr. Kavan is an extremely active consultant in the strategic use of technology, system architecture, and client/server technologies. His main areas of research interest are in interorganizational systems, the delivery of IT services, and technology adoption/diffusion. His work has appeared in such publications as MIS Quarterly, Information Strategy: The Executive's Journal, Auerback's Handbook of IS Management, and Journal of Service Marketing.