Considerable research has been conducted on how consumers evaluate service quality performance. The most commonly accepted approach is the Gap Theory Model which defines service quality as the direction and magnitude of the difference between customers’ expectations and perceptions of the service. This model has been used as an instrument of analysis in several companies. Based on research, SERVQUAL, a multiple item scale for measuring service quality, was developed. The 22-item questionnaire was based on five generic quality dimensions and had become the most popular measure of service quality. Despite its popularity several analysts have suggested that the measure has serious shortcomings that limit its usefulness, such as the overlap of process and outcome, and scaling, and that there may be limitations to widespread use. This article reviews SERVQUAL and the various criticisms of the model in the literature.

**Introduction**

A basic problem of the management of service quality has to do with the nature of services. Although the results of the service, and the instruments it involves may be concrete, the service itself, unlike tangible products, is abstract and difficult to define.

Several definitions of service quality have been suggested in the literature. Some authors have defined service quality as ‘meeting the needs and requirements of customers’ (Murdick, Render & Russell 1990; Smith 1995), while Lewis (1989) states that ‘service quality is how well the service delivered matches the customer’s expectations’. More radical definitions have also been advanced: ‘providing better service than the customer expects’ (Lewis 1989).

In essence, however, the concept of quality has been viewed in the academic world as the difference between expectations and the perceptions of outcomes. The service must correspond to the customers’ expectations and satisfy their needs and demands. Edvardsson, Thomasson & Ovretveit (1994) include in their definition of the ‘right quality’ the various interested parties that need to be considered: ‘The right quality is achieved when expectations are fulfilled, needs satisfied and demands met: those of customers, staff and owners’.

Given the intangible nature of services, it is essential that consumers develop realistic and accurate expectations and that service firms, in turn, deliver these services at or above the level of these expectations. The success of a service firm will depend on how well it meets or exceeds customer expectations. If service quality is to be a cornerstone of a service firm’s strategy, the firm must have the means to measure it. By using distinct measures, firms can identify the most appropriate action and resources can be allocated more efficiently along the service process.

One instrument that was developed to satisfy these goals in service quality measurement is the SERVQUAL scale (Parasuraman, Zeithaml & Berry 1988). In developing this scale the authors defined service quality as the difference between customer expectations for, and perceptions of, actual performance along five dimensions. However, despite initial popularity among practitioners and academics, the model has been criticised on both conceptual, methodological and interpretative grounds.
Objectives and methodology

This article firstly reviews the evolution of SERVQUAL and the theory underlying the model. Since conception by its authors, SERVQUAL has evoked a great deal of response from both academics and practitioners. Some of the applications of the model in different service settings are reviewed and the concerns raised by researchers, as well as other criticisms and shortcomings of the model are highlighted. Various conceptual, methodological and interpretive arguments articulated in the literature are presented both for and against the model. Finally, the article offers suggestions as to the use of the model by practitioners, as well as directions for future research.

The evolution of SERVQUAL

Expected and perceived quality

During the 1980s three American researchers, Berry, Parasuraman & Zeithaml (1985) studied the quality of services. Their early research was about customer-perceived quality in four service industries: banks, credit card companies, stockbrokers, and service companies for household goods. They used focus group interviews with three groups in each industry and expressed the results of their findings as ten factors or dimensions, namely tangibles, reliability, responsiveness, assurance, courtesy, credibility, security, access, communication, and understanding (Parasuraman, Zeithaml & Berry 1985). In a later study (Parasuraman et al. 1988) they reduced the number to five: tangibles, reliability, responsiveness, assurance (which consolidated the competence, credibility, courtesy and security attributes), and empathy (which consolidated the access, communication and understanding attributes).

- **Tangibles** refer to the physical environment in the service organisation: facilities, equipment, staff and their dress, i.e. concrete objects which the customer can easily observe.
- **Reliability** is the company’s ability to perform the promised service. Price agreements and other conditions should be fulfilled, time limits kept and the service performed accurately from the start.
- **Responsiveness** entails performing the services promptly and quickly, helping the customer and being available when he or she needs help.
- **Assurance** covers the knowledge and competence of the staff and their ability to elicit trust and confidence.
- **Empathy** was defined as ‘caring, individual attention the firm provides to its customers’.

The three key points that arise from this research are:

- Service quality is more difficult for customers to evaluate than the quality of tangible goods.
- Customers do not evaluate service quality solely on the outcome of a service; they also consider the process of the delivery.
- The only criteria that count in evaluating service quality are defined by the customer (Zeithaml, Parasuraman & Berry 1990).

One of the most important conclusions in the study by Berry et al. (1985) is that customers’ assessment of service quality is the result of a comparison between their expectations and experience of after service delivery. If their expectations have been met, they are satisfied, if not, they are dissatisfied. If expectations have been exceeded, they are more than satisfied.

In further studies of service quality the three authors found that there are two levels of customers’ expectations of the service: ‘adequate’ and ‘desired’ (Parasuraman, Zeithaml & Berry 1991). The first level is what the customer finds acceptable and the second is what he or she hopes to receive. The distance between the adequate level and the desired level is the ‘zone of tolerance’. This zone expands and contracts and may vary from customer to customer and from one situation to another for the same customer. Similarly they vary, depending on the quality dimension involved.

The Gap Model

The same researchers (Parasuraman et al. 1985) developed a model which depicts how various gaps in the service process may affect the customer’s assessment of the quality of the service. (Refer to Figure 1.) The model is useful in assisting managers and staff to examine their own perceptions of quality and to recognise how much they really understand the perceptions of customers.

**Gap 1** is the difference between the customer’s expectations and management perceptions of customer expectations. Management does not understand how the service should be designed and what support or secondary services the customer requires, i.e. what the right quality for the customer is.

**Gap 2** is the difference between the company’s quality specifications and management perceptions of customer expectations of the service and its quality. Often in an attempt to reduce costs, management places internal restrictions on how a service is to be performed, restrictions which deprive the staff of the opportunity to meet the customer’s expectations.

**Gap 3** is the difference between the quality of the service delivery and quality specifications. Even if the quality of the service is carefully specified in a company, the result in practice may be different from what was intended. Service quality is difficult to standardise since it is often dependent on personal contact between the customer and company staff.

**Gap 4** is the difference between the quality of the service delivery and the quality promised in communicating the product/service. It is important not to promise the customer more than the company can deliver. At the same time, it is important for the company to inform customers about the efforts being made to elevate the quality, which would otherwise not be visible to the customer.

**Gap 5** is the most crucial gap because it indicates the difference between expected and perceived service quality. This gap is a function of the other four gaps:

\[ i.e. \text{Gap 5} = f(\text{gaps 1, 2, 3, 4}) \]

It is this gap that Parasuraman et al. (1985) seek to measure using the SERVQUAL instrument which is discussed below.

The gap model is basically customer-oriented. Quality is realised by the customer after the service has been received and it relates to the difference between expected and perceived quality. The model is also process-oriented because it identifies the gaps that may arise in various parts of the service process, which eventually affect the difference between the customer’s expected and perceived quality. The model is thus based on
what is known as the 'disconfirmation of expectations paradigm' in services marketing literature.

Figure 2 shows a further development of the original gap model. This new model illustrates the interorganisational factors which affect the different gaps. It thereby facilitates an analysis of what caused the gaps and how they can be reduced.

The gap model, which was developed from empirical research, has been used as an instrument of analysis in several companies. Researchers (Parasuraman et al. 1988) have developed an instrument, SERVQUAL, for measuring service quality in the previously mentioned five dimensions (tangibles, reliability, responsiveness, assurance, and empathy).

The SERVQUAL model
The above findings based on Parasuraman et al.'s exploratory research serve as the foundation for the development of SERVQUAL, a multiple item survey tool for measuring service quality. The method was considered by the authors as being 'generic' to all service industries.

There are two parts to the measuring process: the first step is to establish the customers' perception of an ideal service and the second step is to measure the customers' perception of the services provided by a specific company. Perceived service quality is described as the degree and direction of discrepancy between customers' perceptions and expectations. Consequently, SERVQUAL was developed to measure the...
'gap' between expected service and perceived service. This gap corresponds to 'Gap 5' in the gap model described earlier.

The measuring procedure requires the customers to react to 22 statements based on the five quality dimensions. Each of the 22 items was recast into two statements - one requiring respondents to identify which firms in the industry 'should provide'; the other what the customer perceived the firm 'did provide'. There are four or five statements for each quality dimension. (Figure 3 illustrates the five dimensions.)

The respondents are requested to react to the statements on a Likert scale with seven intervals ranging from 'strongly agree' to 'strongly disagree', but with no verbal descriptions for points 2–6. First the expected quality is measured and then the perceived quality. In the first measurement the seven-grade scale produces an 'ideal profile' for each dimension. The profile of customer perceived quality obtained in the second measurement can then be contrasted with the 'ideal'. Deviations between expected quality and perceived quality can then be studied - gap scores or P-E scores could be com-
Service quality measurement: A critical review of the SERVQUAL model

A critical review of the SERVQUAL model

SERVQUAL SURVEY ITEMS

ITEM VARIANCES

Figure 3. Service quality as conceptualised by Parasuraman et al. (1988)

The variables which show the biggest deviations and which the customers see as the most important when assessing quality provide guidelines for quality improvement.

It is important to note that Parasuraman et al. (1988) describe this gap as a measure of service quality as distinct from the measures of satisfaction, on the basis of the nature of the expectations included and the timing involved. In other words, perceived service quality is a 'global judgement or attitude relating to the superiority of the service', whereas measures of satisfaction relate to a service encounter.

The authors further propose that a company's average service quality along each of the five dimensions can then be derived by averaging the SERVQUAL scores across the items on each dimension. Therefore, an overall global service quality score can be obtained by averaging the dimension scores.

To this point the scores are unweighted for relative importance of the different dimensions. Zeithaml et al. (1990) extended the theory to suggest that weighted SERVQUAL scores can also be derived by including importance items that correspond to the original items. The weighted SERVQUAL score is computed as follows:

\[
SERVQUAL \text{ score} = (\text{Perception score} - \text{Expectation score}) \times \text{Importance score}
\]

A revision of SERVQUAL was presented by Parasuraman et al. (1991) which introduced a number of changes. The original SERVQUAL had included negative statements which were subsequently deleted since several researchers experienced problems with these measures (e.g. Carman 1990; Babakus & Mangold 1992). Babakus & Boller (1992) argued that the negatively worded items could be responsible for the factor structure proposed by Parasuraman et al. (1988). A further change focused on the expectations element where respondents were now required to indicate what an 'excellent service would provide' rather than what 'firms in the industry should provide'. Some of the 22 items were changed and/or replaced and minor wording changes were also made.

Uses of SERVQUAL

Parasuraman et al. (1988, 1991) describe SERVQUAL as a concise multiple-item scale with good reliability and validity which offers a number of potential applications across a broad spectrum of services. It provides a basic skeleton through its expectations/perceptions format, encompassing statements for each of the five dimensions. This skeleton, the authors argue, can be adapted and supplemented to fit the needs of a particular organisation.

Parasuraman et al. (1988, 1991) suggest a variety of potential applications of the SERVQUAL model. It can be used periodically to track customer perceptions of service quality relative to that of its competitors. The five-dimensional format of the
model allows a firm to assess its level of service quality along each dimension, as well as overall. The instrument can also be used to categorise a firm's customers into several perceived quality segments (e.g. high, medium, low) on the basis of their individual SERVQUAL scores. These segments can then be compared and contrasted on characteristics such as demographic and psychographic variables so as to gain managerial insights. The instrument can also be used in multi-unit retail companies to track the level of service provided by individual stores and to group the stores into several clusters with varying quality images. An evaluation of store characteristics in the different clusters may reveal attributes that are critical for ensuring high service quality.

In addition, Zeithaml, Parasuraman & Berry (1992) have shown how SERVQUAL can be used in measuring a firm's performance against its competition, which is an important step in the firm's positioning process. Despite SERVQUAL's initial popularity, several studies have, however, identified shortcomings of the model and have failed to provide support for it on a number of issues.

**Applications and shortcomings of SERVQUAL**

**Flexibility vs integrity - Is SERVQUAL a generic measure of service quality?**

The SERVQUAL authors, Parasuraman et al. (1988: 30-31) state: 'It [the modell] provides a basic skeleton... when necessary, can be adapted or supplemented to fit the characteristics or specific research needs of a particular organization'. It should be noted, however, that in 1991, they state: 'Since SERVQUAL is the basic "skeleton" underlying service quality, it should be used in its entirety as much as possible. While minor modifications in the wording of items to adapt them to a specific setting are appropriate, deletion of items could affect the integrity of the scale and cast doubt on whether the reduced scale fully captures service quality'.

A key difference with respect to the SERVQUAL instrument is the extent to which researchers have adhered to the 22-item format. Most researchers, even when using SERVQUAL for its face and/or content validity, have deleted items from or amended the item content so as to make the questionnaire more relevant to a specific service situation. This raises the question as to whether the proposed 22-item scale offers a generic measure.

Similarly a considerable number of researchers have failed to identify the five underlying dimensions. Carman (1990), for example, has identified a greater number of dimensions and others have highlighted the multifaceted nature of services. Some of these are discussed later.

Replication studies of SERVQUAL have, therefore, suggested that the instrument itself requires substantial amendments and extensions to include other key elements affecting the customers' evaluation of the service.

**An extension of the model by Carman**

To address certain shortcomings of SERVQUAL, Carman (1990) proposed an extension of the work by Parasuraman et al. (1988). He replicated the SERVQUAL model in four diverse service industries. The purpose was to investigate six questions related to the SERVQUAL scale:

1. The extent to which the number of dimensions of service quality is generalisable to all settings
2. The robustness of the wording of the SERVQUAL items
3. Service situations with multiple service functions and the role of product quality in bundled retail service offerings
4. The validity of analysing the differences between expectations and perceptions
5. The necessity of administering the expectations battery
6. The relationship between expectations and importance.

Carman (1990) found the dimensions to be 'useful and generally persuasive'. However, certain problems were encountered. Although the model adequately satisfies the first two questions, practitioners 'will need to make some changes in adapting the instrument to a particular setting and will need to make substantial changes in adopting the instrument with respect to questions 3 through 6' (Carman 1990).

Carman's research shows that in using the SERVQUAL model caution should be exercised in reducing the original ten dimensions to five as in any one setting five to seven dimensions were important. These were identified as tangibles, reliability, responsiveness, security, courtesy, personal attention, and trust.

In all settings it would be necessary to alter the wording of some of the individual SERVQUAL items in order to make the item more appropriate to the setting. Some dimensions should have additional items added to those in the original article.

In settings where customers receive a variety of service in a departmentalised way, e.g. department stores and airline travel, customers are able to assess service quality in each of these settings separately. They will evaluate them separately and the overall evaluation of the retailer will be some generalised aggregation of the quality of the parts. In these situations it is necessary to assess service quality in each and then to determine the extent to which each contributes to a perception of overall quality. Carman (1990) suggests that in each of the parts, the same five to seven dimensions are appropriate and can be incorporated into this extended model.

In summary Carman suggests the following:

- The treatment of expectations in traditional applications of the SERVQUAL scale is suspect.
- Importance weights should be included in the measures of service quality.
- The contribution of individual SERVQUAL items to the identified dimensions appears to vary across industries.

The use of SERVQUAL in retail banking

A study conducted by Blanchard & Galloway (1994) sought to determine the perceptions of both customers and staff of the requirements of a quality service in retail banking. The gap model and the SERVQUAL model developed by Parasuraman et al. (1985, 1988) were identified as being the most appropriate for modelling the data, but they found that, although the service gap model provides an excellent basis for analysis, the SERVQUAL model was of limited value.

Blanchard & Galloway argue that the SERVQUAL dimensions are not 'true dimensions' because a fundamental ambiguity lies in the overlap of process and outcome. These are not
separate dimensions, as they are by definition orthogonal and measurable: 'The elements are far too closely interlinked to form the basis of a rigorous analysis of the service situation, and they do not map unambiguously onto the basic classification of outcome, process and expectation' (Blanchard & Galloway 1994).

The researchers contend that although the SERVQUAL dimensions were ostensibly attractive they did not readily align with customer statements of expectations and in many cases customer statements involved at least two of the SERVQUAL dimensions. This interdependence is widely recognised, but as a result the use of the term 'dimension' is invalid, and the lack of clarity it introduces substantially reduces the value of the model.

There are particular problems with reliability which appears to qualify the other four attributes as well as being an independent issue. For example, the requirement for cash to be available at all times in ATMs is clearly a reliability issue, while the politeness is an assurance issue. Reliability is perhaps a prerequisite for quality service in all cases and therefore in a different category from the other dimensions which may or may not be significant in a particular service. The responsiveness dimension too (which contains 'soft' issues e.g. staff listen, as well as 'hard' issues e.g. till open at busy times) is problematic in that it does not focus on specific issues which either the customer or the service designer is likely to address. This would not matter if the dimensions represented some underlying structure, but their ambiguity and overlap suggest that this is not the case.

As a result of these shortcomings Blanchard & Galloway (1994) propose a model based on three dimensions, namely process/outcome, subjective/objective, and soft/hard. These are, according to them, measurable and probably orthogonal. They argue that the model has the advantage of allowing service attributes to be allocated a value within each dimension and provides a classification which avoids overlap and ambiguity. The model also demonstrated that process is far more important than outcome in determining customer perceptions of service quality.

Another study which focused on banks (Brown, Churchill & Peter 1993) also questioned the generic application of the SERVQUAL model. The purpose of their study, however, was to examine the problems associated with using difference scores. However, a major problem arose during their investigation in attempting to modify the wording of the SERVQUAL items to fit the alternative conceptualisation. The authors were concerned by the omission of items they thought would be critical to subjects' evaluation of the quality of service they receive from a bank (e.g. the convenience of a bank's location or its operating hours). They suspected issues such as these were not supplementary or support the original item list because of Parasuraman et al.'s (1988) emphasis on generating a measure applicable across service industries and their consequent focus on items in the analysis that had stable factor loadings across industries. The authors noted that 'it takes more than the simple adaptation of the SERVQUAL items to effectively address service quality in some situations' (Brown et al. 1995).

SERVQUAL in an international recreational service setting
Taylor et al. (1993) tested the applicability or 'generalisability' of the SERVQUAL model using 'confirmatory factor analysis' (Lisrel VII). The study comprised of the following three steps:
1. Confirming the dimensionality and reliability of the original SERVQUAL scale versus importance weighted SERVQUAL.
3. Assessing the influence of service quality on consumer satisfaction.

Their results of step 2 suggest that the summed-and–averaged scales appear adequate measures of service quality in both the original and importance-weighted SERVQUAL scales, while the results of step 3 indicate that service quality evaluations affect positively consumer perceptions of satisfaction across recreational services settings when importance weights are captured and considered.

However, the results of step 1 suggest that neither SERVQUAL nor importance-weighted SERVQUAL's hypothesised five-factor structures are confirmed by confirmatory factor analysis in the research settings. Furthermore, evidence is presented suggesting that the reliability of the alternative scales is not significantly different, i.e. the addition of importance measures does not appear to enhance the reliability of the scale.

These results lead to the conclusion that the SERVQUAL scale appears as a 'doubtful' scale for measuring service quality in the leisure activities setting and that although the SERVQUAL scale appears to have some support, there may be several limitations that must be addressed prior to widespread use in international applications.

SERVQUAL in a retail setting
An examination of the usefulness of SERVQUAL in a retail setting was conducted by Finn & Lamb (1991). The researchers posited that if the SERVQUAL scales possess construct validity (i.e. if the 22 items in the model measure the five dimensions) in a retail setting, then a survey of retail store customers should produce results that conform to the model.

Using Lisrel and the chi-square goodness-of-fit statistical technique to their results indicate that the SERVQUAL measurement is not appropriate in a retail store setting and challenge the validity of SERVQUAL scales as measures of the determinants of perceived quality in a retail setting. Finn & Lamb (1991) attribute the failure of the model to the following:

- The SERVQUAL scales do not capture the essence of the service quality construct in retailing. Retailing does not fit into the data set of the original four industries identified by Parasuraman et al. (1988), i.e. banking, credit card, repair and maintenance, and long distance telephone companies, and the scales are therefore inappropriate for measuring the five dimensions.
- Perceived quality in retailing is not a function of the five constructs identified by Parasuraman et al. (1988).

The abovementioned service categories are polarised closer to the pure-service end of the pure-service/pure-goods continuum than is retailing. It may well be that consumers use different criteria to evaluate competing goods retailers than they use to evaluate retailers that are exclusively service firms.

This study raises the question of whether the five dimensions used in the model are generic, and also whether the
model measures the determinants of perceived service quality in all service industries. The results reported by Finn & Lamb (1991) suggest that the construct validity of SERVQUAL should be examined on an industry by industry basis before it is used to gather consumers' perceptions of service quality.

A similar study of retail apparel customers' expectations and perceptions of service quality offered in retail specialty stores was conducted by Gagliano & Hathcote (1994). Using the SERVQUAL model the researchers discovered that they had to reanalyse their data using a four dimensional factor analysis instead of the original five due to the overlapping of two of Parasuraman's factors and the low ranking of a fifth factor. The five determinants did not factor out as expected. The researchers replaced responsiveness, assurance and empathy with personal attention and convenience arguing that these were more appropriate in the apparel specialty setting.

Gagliano & Hathcote (1994) suggested that the SERVQUAL scale should be used cautiously and that it should be refined before it can be accepted as a valid measurement scale in apparel specialty store settings.

Other criticisms of SERVQUAL

The use of difference-scores

While not generally recognised, the conceptualisation of service quality as a difference-score leads to a number of potential problems. These problems were reviewed and an investigation was carried out by Brown et al. (1993) to determine if they arose empirically with SERVQUAL. In an earlier publication by Peter, Churchill & Brown (1993) the authors cautioned the use of difference-scores as measures of constructs. They concluded that difference-scores are:

- less reliable than other measures
- may appear to demonstrate 'discriminant validity' (involves the extent to which a measure is novel and does not simply reflect some other variable)
- may only be 'spuriously' correlated to other measures since they typically do not discriminate from at least one of their components
- may exhibit variance restriction (which occurs when one of the component scores used to calculate the difference-score is consistently higher than the other component).

Brown et al. (1993) also explored a nondifference-score conceptualisation of the same facets of service used in the SERVQUAL measure. The investigation indicated that the problems with SERVQUAL, brought on by its measurement as a difference-score, manifest themselves empirically. Although SERVQUAL had high reliability, its reliability was below that of a nondifference-score measure of service quality. Moreover, not only did SERVQUAL fail to achieve discriminant validity from its components, but the perceptions component, by itself, performed as well as the difference-score on a number of criteria. SERVQUAL also exhibited variance restriction effects and the distribution of SERVQUAL scores was nonnormal.

Brown et al. (1993) established that the nondifference-score measure did not exhibit these problems. Moreover, it displayed better discriminant and nomological validity properties. They claimed that the nondifference-score measure performed better than SERVQUAL on a number of important psychometric and statistical considerations. It did so while requiring subjects to respond to only half as many items (22 instead of 44), and thus is twice as efficient. The nondifference-score measure also allowed subjects to compare their expectations and perceptions directly and did not restrict them to some arbitrary, linear difference. Thus, the authors state:

> These disappointing results raise serious doubts about the correspondence between the SERVQUAL measure and the theory underlying it. It seems the theory is incorrect in specifying five components of service quality or that the measure is incorrect in capturing only one component of service quality when theory suggests there are five dimensions. In future research it should be investigated whether SERVQUAL is only assessing a unidimensional construct, rather than the question of the validity of the theory, or measure, or both (Brown et al. 1993).

Cronin & Taylor (1992) also found that their measure of service performance produced better results than SERVQUAL. Their nondifference-score measure consisted of the perception items used to calculate SERVQUAL scores. This measure assessed service quality without relying on the disconfirmation paradigm. Brown et al. (1993) have presented a new measure that performs as well as the perceptions component of SERVQUAL yet includes a comparison of perceptions with expectations.

The results of a qualitative assessment by Teas (1993) of the original SERVQUAL model (1988) and the revised model (1990) also indicate that the measures lack discriminant validity with respect to the concepts of ‘attribute importance’, ‘performance forecasts’, and ‘classic attribute ideal points’. This suggests that a considerable portion of the variance in the SERVQUAL expectations measures may be caused by respondents' misinterpretations of the question rather than to different attitudes or perceptions.

Parasuraman et al. (1993, 1994) defended the above criticism by arguing that the deficiencies of the difference-score conceptualisation are not as severe as they are made out to be: 'The superior predictive power of the P-only (perceptions) measure must be balanced against its inferior diagnostic value'. Furthermore they argue that the difference-score formulation 'provides richer, more accurate diagnostics for improving service quality' (1993) and that managers can continue to have confidence in the difference-score conceptualisation of SERVQUAL. The use of difference scores is nevertheless questionable and further research is deemed necessary.

Practical issues and the timing of administering SERVQUAL

The major shortcoming of the SERVQUAL model as identified by Carman (1990) is that Parasuraman et al. (1988) suggest collecting data on consumers' expectations of the service they are about to receive, presumably as they come in the door; then to ask a very similar battery of questions on consumers' perceptions of the service received, as they leave. Then the authors recommend finding the difference between the perceptions and expectations and using this value in the quantitative analysis. Carman finds that this procedure is not very practical nor is it the best analytical procedure and suggests alternatives on how to get around this problem.

In addition to the practical difficulties identified by Carman (1990) in administering the SERVQUAL model, timing difficulties in the administration of the two sets of SERVQUAL statements have also been experienced by other researchers. It
would appear from Zeithaml et al.’s work (1990) that the two batteries of questions should be administered at the same time and not be related to an encounter. Two studies emphasising the importance of the timing are briefly discussed below.

Bolton & Drew (1991a) offer an examination of the conceptualisation and operationalisation of service quality which questions the model of Parasuraman et al. (1985, 1988). They developed a model of the longitudinal effect of a service change on perceptions of service quality. Unlike the previous research of Carman (1990), Bolton & Drew’s analysis investigates temporal changes in individual attitudes. That is, the literature on service quality typically measures the construct and its underlying dimensions using cross-sectional data rather than attempting to measure attitude changes in service quality perceptions over time. Bolton & Drew criticise the cross-sectional approach for failing to account for the possibility that the factors which explain differences among consumers’ attitudes at a given time \( t \), may not be the same as the factors that cause changes in a given consumer’s attitude at the time \( t+1 \). Thus, the magnitude and direction of the gap between customer expectations prior to a service and the evaluation of the service received, impact on the level of consumer satisfaction/dissatisfaction.

The findings of another study conducted by Clow & Vorhies (1993) is adequately summarised below:

...the simultaneous measurement of consumer expectations and evaluation of service quality led to biased measures of expectations. Expectations continue to play an important role following the consumption experience. For dissatisfied consumers, the gap between expectations and experience gets larger. For satisfied consumers, the gap becomes smaller. For accurate measures of service quality, consumer expectations should be measured before the service experience and evaluation of the service after the patronage occurs (Clow & Vorhies 1993).

The exclusion of price or value

The SERVQUAL model has also been criticised for ignoring price or value. Consumers’ expectations and consequent evaluation of the service must be affected by price. Zeithaml et al. (1990) do argue that a key influence of customers’ expectations is price, but the conceptualisation of quality as distinct from value is why it is not perpetrated into the definition of perceived quality. Smith (1995) argues that this distinction is of little value when assessing consumer evaluations of a firm or its competitors: ‘Expectations of an excellent organisation would be of little value to those targeting the low-price/reasonable quality segment’. As a result, Freeman & Dart (1995) included fees as a dimension in their study of accounting firms.

Criticisms of the expectations scale

Two key issues have arisen as a result of the practice of calculating the gap between consumers’ expectations and perceptions as a measure of service quality. Firstly, Teas (1993, 1994) questions the meaning of the expectations measure and suggests that a substantial portion of the variance in the expectations scale is due to differences in respondents’ interpretations of the question being asked rather than to the variance in respondents’ attitudes. In a study conducted by Smith (1995) the revised expectations measure (i.e. from ‘should’ to ‘excellent companies will’) appeared to have little advantage over the original scale. The mean score for the expectations scale was 6.401. She noted that of the 29 items the lowest mean score was 5.13. These high scores for the expectations scale are likely to result in negative P-E scores, which affects both the diagnostic utility of the measure and the underlying conceptual interpretation (Smith 1995).

The second issue is whether the expectations battery should be administered at all. Smith (1995) noted that several researchers neglected to measure expectations and several others highlighted the independent effects of perceptions on consumer evaluations of satisfaction or quality (Carman 1990; Bolton & Drew 1991a; Cronin & Taylor 1992). Consequently, the usefulness of the adoption of the disconfirmation paradigm was brought into question.

Cronin & Taylor (1992, 1994) questioned the five dimensions of the SERVQUAL model by arguing that the disconfirmation-based paradigm of the model is flawed. They also provided empirical evidence that service quality should be measured as an attitude. The authors have proposed an alternative measure called SERVPERF, which is a performance-based measure of service. Apart from other literature supporting the performance-based paradigm (Babakus & Boller 1992; Babakus & Mangold 1992; Boulding et al. 1993), Cronin & Taylor (1994) cite the following as the ‘most telling evidence thus far’:

Our results are incompatible with both the one-dimensional view of expectations and the gap formation of service quality. Instead, we find that service quality is directly influenced only by perceptions of performance (Boulding et al. 1993).

Consumers’ interpretation of gap scores

Parasuraman et al.’s (1988, 1991) conceptual interpretation of gap scores suggest that where gap scores are positive the respondent perceives higher quality and therefore would offer more favourable evaluations.

The findings of a study conducted by Smith (1995) illustrate how the respondent’s evaluation of the standard of a service does not solely derive from a comparison of expectations of excellence with perceived performance, but also from other factors which may include the importance of elements of a service, and experience of alternative suppliers. Furthermore, attribution and halo effects may cause respondents to perceive their chosen supplier more favourably. The impact of response sets due to the nature of the scales and the respondents, should also be considered.

Midpoint of the perceptions scale

One further reason why SERVQUAL may present meaningless information with respect to the interpretation of gap scores is where the respondent ‘does not know’ and may therefore record a ‘4’ on the perceptions scale. Whether expectations are positive or negative, the resultant score would suggest inappropriate action to the practitioner (Smith 1995).

Smith (1995) identified at least five broad meanings which respondents assigned to the midpoint (‘4’) of the scale, i.e.:

1. An evaluative response including intention
2. Observations of lack of consistency of the service provider
An appropriate measuring system will help service managers ensuring system in service industries is immense. Managers are often forced to use quality measures which are appropriate for rating product quality but not service quality, or that lack both conceptual bases and empirical generalis ability. This can result in faulty analyses which could lead to poor decision-making. Thus, service firms are still faced with uncertainty when trying to identify an appropriate measure of service quality. Given the importance of measuring and controlling service quality and the shortcomings of existing efforts, additional research in this area seems warranted.

Managerial implications and recommendations

The results reported by various researchers suggest that the construct validity of SERVQUAL should be examined on an industry by industry basis before it is used to gather consumers' perceptions of service quality. Managers are advised to carefully consider which issues are important to service quality in their specific environments and to modify the SERVQUAL scale as needed. The nondifference-score version of the scale can serve as a useful starting point for these modifications.

Cronin & Taylor (1992, 1994) have suggested a performance-based measure of service as an improved means of measuring the service quality construct. They have consistently argued that managers should not include consumer expectations in measures of service quality, although expectations can impart valuable information 'if their unique effect on purchase behaviours and performance perceptions are conceptualized properly'. In addition to performance-based measures, performance-based maps would be of benefit.

Other researchers have provided a means of overcoming psychometric problems with SERVQUAL. Brown et al. (1993) and Carman (1990) suggested that statements be rephrased and that respondents record their evaluation on a scale ranging from 'much worse than I expected' through 'neutral' and 'much better than I expected'.

Other issues listed in this article also need to be carefully considered before managers apply SERVQUAL to a particular setting. The continuing debate about SERVQUAL is encouraging for managers, as a universally acceptable tool should begin to emerge once the conceptual, methodological and interpretative issues surrounding SERVQUAL are resolved.

Presently, however, an important implication is that managers should not treat SERVQUAL as an 'off the shelf' (Finn & Lamb 1991) measure of perceived quality. Considerable refinement is needed for specific companies and industries before applying the model.

Future research

The criticisms cited in this article should not serve to render the SERVQUAL model redundant. However, it should encourage researchers to use the existing model as a basis for further exploration.

The need for a theoretically sound and generalisable measuring system in service industries is immense. Managers are often forced to use quality measures which are appropriate for rating product quality but not service quality, or that lack both conceptual bases and empirical generalisability. This can result in faulty analyses which could lead to poor decision-making. An appropriate measuring system will help service managers to develop quality standards which more accurately represent the activities that result in the provision of the service.

Despite the various methods suggested with respect to measuring service quality, none except SERVQUAL has received extensive empirical testing. The debate about SERVQUAL makes it clear that the conceptual clarity about the dimensions of service quality has not as yet been achieved. To have an impact on service strategy, the studies of service quality will need to incorporate variables other than those identified in SERVQUAL, e.g. price/quality or value relationship. The incorporation of these variables would help to clarify the relative priorities of not only the dimensions of quality, but also the other 'service winners' (Rosen & Karwan 1994).

Conclusion

As service quality has become increasingly important to service strategies, its assessment has become more critical; strides have been made in recent years to measure it. The most popular measure, SERVQUAL, involves the subtraction of subjects' expectations of the service they would receive from their perceptions of the service they actually did get with respect to specific items. The differences are averaged to produce a total score for service quality. While the scale attempted to provide a generalisable measure of service quality, a number of studies have shown that such a claim may be inappropriate. In addition, despite initial popularity among practitioners and academics, SERVQUAL has been criticised on both conceptual and methodological grounds.

Several researchers have questioned the extent to which the model is generic to all industries, and similarly, have failed to identify its five underlying dimensions. Other issues vital to the consumer's evaluation of the service are not addressed by the model. Some researchers have also questioned the relevance of expectations in the model as well as the timing and frequency of administration. Others have argued against the use of difference-scores which has led to psychometric problems in the model.

The P-E formulation has been shown to be problematic in that the scores are usually negative, which questions the current base employed by SERVQUAL, and the reasons why positive scores might be obtained. The use of a midpoint scale in the model also has its shortcomings.

The SERVQUAL model remains an issue of debate in contemporary services marketing literature. Some researchers have recommended the abandonment of the model altogether, while others continue to modify or extend the model in their applications. Despite its criticisms, the ground-breaking work of Zeithaml et al. (1990) has paved the way for a deeper understanding of service quality and the eventual development of an important measurement tool.

References


