

# Measuring internal service quality: Comparing the gap-based and perceptions-only approaches

Dr Alistair Brandon-Jones<sup>1</sup>

*Lecturer in Operations and Supply, Bath School of Management, University of*

*Bath, Bath, BA2 7AY. UK*

*Tel: +44 (0) 1225 383 886*

*E-mail: [a.brandon-jones@bath.ac.uk](mailto:a.brandon-jones@bath.ac.uk)*

Dr Rhian Silvestro

*Associate Professor in Operations Management, Warwick Business School,*

*University of Warwick, Coventry, CV4 7AL. UK*

*Tel: +44 (0) 2476 522 991*

*E-mail: [rhian.silvestro@wbs.ac.uk](mailto:rhian.silvestro@wbs.ac.uk)*

*January, 2010*

*Submitted for review by*

*The International Journal of Operations and Production Management*

*Special Issue of the EUROMA 15th International Annual Conference, Groningen,*

*2008*

---

<sup>1</sup> The authors wish to thank all those who participated in this research project. In addition, they would like to express their gratitude to the editors and reviewers for the time and effort they gave in reviewing this paper. The feedback provided was extremely useful in improving the work.

# Measuring internal service quality: Testing two approaches

## **Abstract**

**Purpose** – This paper builds upon the debate in the service quality literature regarding both the theoretical and practical effectiveness of expectations data in the measurement of internal service quality. Gap-based and perceptions-only approaches to measuring internal service quality are tested and their respective benefits and limitations evaluated.

**Design/methodology/approach** – The internal service context used in this study is the provision of e-procurement software, training, and user support in four organisations. The two approaches are evaluated in terms of reliability and validity, as well as pragmatic aspects of survey administration.

**Findings** – The various tests carried out indicate that both the gap-measure and perceptions-only measure are reliable and valid, the latter being the marginally higher performer. Both approaches were found to have benefits and limitations, and so the empirical study, combined with contributions from the literature, generates some understanding of the internal service context in which the two approaches might be appropriate.

**Research limitations** – This research study was confined to a particular type of internal service context: an internal e-procurement service. There is a need to further test alternative measurement approaches in different internal service contexts in order to further refine understanding of internal service quality measurement.

**Practical implications** – For operations managers, the paper clarifies the basis on which they might choose between the two approaches to internal service quality measurement.

**Originality/value of the paper** – This study is the first to directly test and compare the relative merits of these two approaches to internal service quality measurement. The paper also offers insights as to the operational contexts in which each approach might be appropriate.

**Key words** – Internal service quality, Service quality measurement, SERVQUAL, e-procurement

**Paper type** – Research paper

## **Introduction**

The management of internal service quality can be traced back to Ishikawa's concept of the 'voice of the customer' (1985) and has been an emerging theme in the service operations and marketing literature over the past two decades (George, 1990; Davis, 1991; Stauss, 1995; Ahmed and Rafiq, 2000). Internal service quality is defined as the perceived quality of service provided by distinctive organisational units or the people working in these, to other units or employees within the organisation (Stauss, 1995). Internal services create a network of functional units which are linked together with the aim of delivering service to external customers (Marshall *et al.* 1998). As such,

delivering service quality to external customers hinges on delivering service quality across internal supply networks.

Johnston (1999, 2005) argues that many of the contributions to the literature on internal service quality emanate from services marketing, and that there is a need for the operations management discipline to contribute to the development of frameworks and tools for improving the management of internal customer relationships and networks. Within this research agenda, the measurement of internal service quality is key, as it provides a basis for continuous improvement (Koska, 1992; Reynoso and Moores, 1995; Young and Varble, 1997; Frost and Kumar, 2000), and helps to enrich traditional cost-based approaches of supplier performance assessment (Large and König, 2009).

Within the external service literature, SERVQUAL (Parasuraman *et al.* 1988) has been at the centre of the debate as to how service quality should be modelled and operationalised into effective measurement systems (Buttle, 1996; Silvestro, 2005). Much of this discussion has focused on whether the construct should be based on the gap between expectations and perceptions, or whether perceptions-only measures of service quality might be more reliable and effective. More recently, these issues have been debated with respect to internal service quality, with questions raised over the transferability of external service quality measures to internal services (Reynoso and Moores, 1995; Frost and Kumar, 2000). There remains a need to compare and evaluate gap-based and perceptions-only measures of perceived quality in this context.

This paper reports the findings of a study which tests the two approaches both in terms of their theoretical underpinnings and also in the light of practical considerations regarding the design of measurement systems. The paper takes an operations management rather than a marketing perspective, in that the measurement instruments are evaluated as tools for identifying operational improvement priorities with a view to improving process design and delivery. The internal service context used in this study is the provision of e-procurement software, training, and user support. Analysis is based on survey data from 274 internal customers of e-procurement services provided by the procurement departments in four organisations. The alternative measures are evaluated in terms of reliability, content validity, construct validity, and predictive validity, as well as practical considerations concerning implementation.

We begin by reviewing the debate surrounding alternative approaches to measuring internal service quality. This gives rise to research questions which call for an evaluation of gap-based and perceptions-only measures in internal service contexts. The methodology for our study is then described in detail, followed by analysis of the two alternative measures of internal service quality. The discussion section examines our research questions in light of our analysis, presents limitations, and identifies opportunities for further work. Finally, conclusions are drawn based on this study.

### **Literature review**

The notion of internal customers originates from TQM's 'next-operation-as-customer' perspective (Ishikawa, 1985; Deming, 1986; Juran, 1989; Ratcliffe-Smith and Brooks, 1993), in which organisations can be viewed as a network of functional units, linked together with the aim of delivering service to external customers (Marshall *et al.* 1998). Each unit receives inputs, transforms them, and delivers the output to the next operation in the chain – their internal customer. Each link in the chain represents an interaction between internal service providers and internal customers (Finn *et al.* 1996). Whilst the internal customer concept found in TQM literature shares some similarities with internal marketing, the key difference is that internal marketing largely focuses on how the company serves its internal customers (Marshall *et al.* 1998). In contrast, the next-operation-as-customer perspective adopted in this study usually views the service provider as an organisational unit or even an individual (Heskett *et al.* 1994).

#### *Measuring internal service quality*

Compared with external service research, there is relatively limited research focused on internal service quality measurement. This is partly a consequence of the marketing background of many service quality academics (Iacobucci *et al.* 1994) and the multi-disciplinary nature of internal service (Hallowell *et al.* 1996; Farner *et al.* 2001).

Attempts to measure internal service quality follow two common approaches. The first is to adopt a gap-based measure of internal service quality, usually through the application of the SERVQUAL scale (Parasuraman *et al.* 1988). These applications range from almost exact replication (Young and Varble, 1997; Auty and Long, 1999;

Kang *et al.* 2002), to minor changes (Chaston, 1994; Hill and McCrory, 1997; Frost and Kumar, 2000), to addition and deletion of dimensions (Kuei, 1999; Large and König, 2009), through to major departures from the scale (Boshoff and Mels, 1995; Reynoso and Moores, 1995; Brooks *et al.* 1999; Stanley and Wisner, 2001). The second approach has been for researchers to develop perceptions-only measures of internal service quality, usually from scratch. These include the provision of banking services (Lewis and Gabrielson, 1995), insurance services (Hallowell *et al.* 1996), procurement services (Cavinato, 1987; Hendrick and Ruch, 1988; Rossler and Hirsz, 1996; Finn *et al.* 1996; Croom and Brandon-Jones, 2007), and generic internal services (McDermott and Emerson, 1991; Gilbert, 2000; Bruhn, 2003).

Before exploring the debate concerning alternative approaches to internal service quality measurement, let us first consider the differences between external and internal customers which have led some academics to call into question the transferability of external service quality measurement approaches to internal services.

#### *Differences between internal and external customers*

There are some well documented characteristics of internal customers which are likely to affect the measurement of internal service quality and which may challenge the transferability of approaches developed to measure the service perceptions of external customers. The key differences concern customer choice and expertise (Stauss, 1995; Marshall *et al.* 1998; Farner *et al.* 2001; Bruhn, 2003).

External customers can typically choose where to take their business (Finn *et al.* 1996) and have the option of exiting unsatisfactory relationships. Such free market forces motivate organisations to provide excellent service quality in order to retain customers. By contrast, internal suppliers have tended to occupy a monopolistic position, with internal customers often given little choice over their service provider regardless of quality or cost (Gremler *et al.* 1994; Auty and Long, 1999; Farner *et al.* 2001). Therefore, whilst repeat custom is a sign of good service in external settings, internal customers may keep coming back simply because they have no alternative (Albrecht and Bradford, 1990).

A further important difference between external and internal customers concerns the way they evaluate quality. Many external service quality measures are largely based on experience properties of service quality, because, it is argued, services have

few search properties and it is often difficult to assess credence properties (Parasuraman *et al.* 1985). However, Marshall *et al.* (1998) state that because internal customers are ‘professional’ consumers of internal services, they are far more knowledgeable than most external customers with regard to service provision. As such, they may be in a stronger position to assess credence properties, such as, for example, the competence of service providers. This view is borne out by a number of internal SERVQUAL applications that have omitted the tangibles dimension when measuring internal service quality (Brooks *et al.* 1999; Heskett *et al.* 1997; Kuei, 1999; Large and König, 2009). Unlike external customers, who may be impressed with cosmetic features, internal customers may see these same elements as excessive and wasteful (Paraskevas, 2001). Furthermore, there is often little face-to-face interaction between internal customers and internal suppliers (Young and Varble, 1997). As a result, tangible elements such as physical layout, equipment and clothing, may be of little concern when making service quality assessments of internal suppliers.

Finally, the knowledge and experience of internal customers may mean that they are less influenced by high-expectations social norms found in external service research. For example, in a recent study of internal service quality, Large and König (2009) report expectation levels which are lower than many reported external service quality expectations. Recognition of the differences between external and internal customers has led a number of researchers to question the transferability of service quality measurement approaches developed for external customers to internal customer contexts.

#### *The alternative internal service quality measures*

Many internal service quality measures are based on the disconfirmation paradigm, which states that service quality is determined by the gap between expectations and perceptions of performance. Whilst this perspective is dominant within the service literature, concerns remain over its theoretical applicability. Firstly, there are objections to defining a construct as the difference between two other constructs – expectations and perceptions (Carman, 1990; Teas, 1993, 1994; Brady *et al.* 2002). Secondly, there is the argument that disconfirmation theory is more appropriate when measuring the transaction-specific concept of customer satisfaction (Cronin and Taylor, 1992, 1994). Finally, the gap-approach can lead to a ‘service paradox’,

whereby simply lowering customer expectations has the effect of ‘increasing’ service quality, because the gap between expectations and perceptions is reduced (Grönroos, 1988). Based on these theoretical concerns, a number of authors argue that a perceptions-only (i.e. direct / non-difference) approach is more appropriate in measuring perceptions of service quality (Cronin and Taylor, 1992; Smith, 1995; Van Dyke *et al.* 1997). For example, Cronin and Taylor (1992, 1994) propose a performance-only measure of service quality. SERVPERF uses the same 22 perception items as SERVQUAL, but does not include the set of expectations statements.

Babakus and Boller (1992) suggest that whilst service quality measurement based on perception-expectation gaps is intuitively appealing, “difference scores do not provide any additional information beyond that already contained in the perceptions component of the SERVQUAL scale” (pp.263). Parasuraman *et al.* (1994a) accept that performance-only measures of service quality tend to have higher predictive accuracy, but this comes at the cost of diagnostic value: “SERVQUAL could be superior in terms of pinpointing areas of deficiency within a company” (pp.116). Dean (1999) concurs with this view and supports the use of gap scores because of their diagnostic value. As an example, if a customer rates expectations of two service attributes at 5 and 7 respectively, and perception of these two attributes at 4 and 5, a manager using a perceptions-measure would conclude that the first attribute is the key problem area, even though the gap between expectations and perceptions is much higher for the second attribute. Furthermore, direct measures of service quality may suffer from over-inflation of customer service ratings (Peterson and Wilson, 1992).

A number of authors have noted that expectations scores are misleading because the most likely response to statements on expectations of service delivery is ‘strongly agree’ (Carman, 1990). Individuals are often driven by the ‘I-have-high-expectations’ social norm and this creates a bias towards social desirability (Brandon-Jones *et al.* 2010). Social desirability is a form of common method bias (Phillips and Clancy, 1972; Podsakoff *et al.* 2003) that arises from the tendency of some individuals to inflate responses in line with what is regarded as socially acceptable, referred to by Howard *et al.* (2007) as the ‘bandwagon effect’. If expectations scores are consistently high, perceptions will be the dominant contributor to gap scores. However, particularly in an internal service context, knowledge and experience may have an effect on the level of expectations and may be less influenced by high-

expectations social norms found in external service research. For example, in a recent application of the gap-based measure of internal service quality, Large and König (2009) report expectations scores ranging from 4.8 to 6.4 (p28), averaging 5.96 (on a 1-7 scale). Not only are these scores lower than those reported in many studies of external customer expectations, but they also exhibit enough variation to be of practical use.

Considering approaches to data collection, there are some concerns as to when internal customers are asked about their expectations. Clow and Vorhies (1993) argue that post-service expectations scores are strongly influenced by customer perceptions of services. Customers who are happy with the service tend to understate expectations, whilst dissatisfied customers will tend to overstate them. As such, the collection of expectations after the event creates risks to data reliability.

Finally, the gap-approach may suffer from the boredom factor of two administrations, one for expectations and the other for perceptions (Bouman and Van der Wiele, 1992). Reynoso and Moores (1995) have proposed an alternative approach to measuring internal service quality with the intention of obviating the practical problems of administering lengthy two-part questionnaires, whilst retaining a gap perspective. They advocate surveys based on single statements which capture the perceptions-expectations gap rather than simply using the perceptions half of paired statements (Table I provides examples of item formulation). Testing the scale, they conclude that it combines the benefits of the academic grounding in disconfirmation theory with desirable economies in questionnaire length.

**Table I. Example of single-item gap-based measures (Reynoso and Moores, 1995)**

Quality Factor	Well below my expectations			Well above my expectations		
	1	2	3	4	5	6
Availability of support to deal with queries	1	2	3	4	5	6
Speed of response to user queries	1	2	3	4	5	6

However, the single-item gap-based approach should not necessarily be regarded as the solution to internal service quality measurement. Whilst it combines some of the advantages of both methods, it is also vulnerable to the disadvantages of both, suffering particularly from a lack of conceptual lucidity. The data resulting from such a survey do not provide the insights into expectations that are characteristic of the gap approach; but neither does the measure have the simplicity and clear meaning of the

perceptions-only approach. Indeed the benefits of gap-based measures in terms of diagnostic value are lost; whilst the problems associated with interpretation of the perceptions-only measure may be amplified using the single-gap measure approach.

### **Research objectives**

Within the internal service quality literature, there remains a need to assess the psychometric and practical value of scales based on the gap-approach as opposed to the perceptions-only approach. Therefore, the main research objective of this study is to compare two internal service scales in terms of reliability, validity, and pragmatic aspects of survey administration. The study focuses on the following questions:

- *How reliable and valid is a gap-based measure of internal service quality?*
- *How reliable and valid is a perceptions-only measure of internal service quality?*
- *What are the benefits and limitations of each approach?*

A survey of internal service quality was conducted in order to conduct this evaluation. There now follows an explanation of the survey design and of the approach taken to collect and analyse the data.

### **Research design**

In order to evaluate the relative merits of the gap-based and perceptions-only approaches to internal service quality measurement, a theoretical sample of internal services were invited to participate in the survey (Eisenhardt, 1989). The involvement of multiple service sites, rather than a single service would provide a more robust basis for testing the measurement instrument. However in order for the survey instrument to be effective in measuring expectations and perceptions, it was necessary to survey the customers of similar services which could be evaluated on the same criteria. To ensure comparability of the internal services, the internal purchasing departments of four UK organisations were chosen, all providing e-procurement services to their internal customers.

The four departments purchased their e-procurement system from an external software supplier, and then customised the software for internal application. This software supported purchase ordering, authorisation, receipting, invoicing, payment, and reporting. The purchasing departments were responsible for training internal customers across their organisations in the use of the software, as well as providing ongoing user support. The fact that the four organisations used the same software package affords some comparability between the organisations, in that differences in their expectations and perceptions could not be explained in terms of their use of different software, but was rather based specifically on their expectations and perceptions of the internal service received from their e-procurement departments. The four organisations covered a broad range of procurement activity, ranging from the procurement of high value, bespoke services to the purchase of low value, high volume commodities. They also varied in terms of size, budget, number of suppliers, number of internal customers, and level of e-procurement integration (Table II). The internal customers in the study manifested the characteristics which typically distinguish internal customers from external customers as discussed earlier. They were locked into the service and were expected to use it regardless of their satisfaction. They had relatively little face-to-face contact with the purchasing department and tangible elements of the service (other than characteristics relating to the software) were of low priority. Finally, they were experienced, ‘professional’ users of the service and their expectations would have been shaped by past experience as well as credence properties. Therefore, the context selected was considered to be appropriate for testing the two internal service quality measurement approaches.

**Table II. General characteristics of the four organisations**

	<b>Org 1</b>	<b>Org 2</b>	<b>Org 3</b>	<b>Org 4</b>
<b>Number of employees (FTE)</b>	26,500	800	200	450
<b>Yearly budget (total)</b>	£1.6 billion	£45 million	£18 million	£40 million
<b>Yearly budget (goods and services)</b>	£600 million	£16 million	£6 million	£15 million
<b>Requisitions per annum</b>	150,000	4000	2000	2900
<b>Active suppliers</b>	13,000	2500	800	2300
<b>E-procurement service users</b>	156	44	41	54
<b>Level of financial systems integration</b>	Extensive	Limited	None	Limited

### *Questionnaire design*

Previous internal service research has established the need to identify the quality factors pertinent to particular contexts, and to build these into survey questionnaires, rather than simply replicating existing scales such as SERVQUAL (cf. Boshoff and Mels, 1995; Kuei, 1999; Brooks *et al.* 1999; Stanley and Wisner, 2001, Large and König, 2009). Therefore, a 33-item measure of internal service quality in an e-procurement context was developed from scratch. This drew on external service, internal service, information systems, and e-service literature, in addition to semi-structured interviews with e-procurement service providers and internal customers. Full details of this scale development are documented in Brandon-Jones (2006 and 2008). The survey consisted of paired-statements relating to different aspects of internal service quality (see appendix 1 for item details and definitions). The first set of statements related to expectations and the second to perceptions, both with 1-7 Likert scales from ‘strongly disagree’ to ‘strongly agree’. In addition, there was a single question asking users to rate the overall quality of e-procurement service provision – the *overall e-procurement quality rating (OEPQ)* – anchored on a 1-7 Likert scale from ‘very poor’ to ‘excellent’. Having a separate independent measure of internal service quality enabled an evaluation of the relative validities of the gap-measure and perceptions (Parasuraman *et al.* 1988; Pitt *et al.* 1997). There is a good deal of support for the use of single-items scales in measuring psychological constructs (cf. Sackett and Larson, 1990; Scarpello and Campbell, 1983; Wanous *et al.* 2007). Wanous *et al.* (2007) argue that the additional space required for multi-item construct measures is often impractical and can damage response rates. In addition, there may be face validity concerns if respondents feel they are being asked repetitious questions. This last point was important, as we were essentially looking to measure the same construct, internal service quality, twice – once with the 33 internal service quality items and once with the single-item *OEPQ rating*. Finally, single-item measures were used for control variables – organisational size, IT skills rating, and purchasing experience.

Initially, academic colleagues with expertise in the service quality measurement and e-procurement literatures were asked for feedback on the survey questions, structure and format. Subsequently, 18 e-procurement users in two organisations not involved in the survey were sent the proposed questionnaire and all returned annotations commenting on its clarity and ease of use. The academic and practitioner

feedback that was received helped to refine question wording, although no major changes were required.

#### *Data collection and preparation for analysis*

Within this study, the population was defined as all internal customers of the e-procurement software and support provided by the purchasing departments across the study organisations. As there were only 295 eligible e-procurement users within the population frame, a census (100% sample) was applied (Easterby-Smith *et al.* 1997). To encourage buy-in to the survey process and secure a high response rate, contact was made with all potential respondents prior to sending out questionnaires, to explain the purpose of the research and invite their participation. This was the first survey to have been implemented by the purchasing departments, so staff were not survey weary and were fortunately willing to engage with the research process. Consequently, 274 usable questionnaires were returned, representing an extremely high response rate of 92.9%.

Data were entered in SPSS 14.0 for statistical analysis. Appendix 1 provides details of means and standard deviations for expectations, perceptions, and gap scores. Considering non-response bias, no significant differences were found between the means of early and late respondents for any variables. T-tests and an overall test of randomness found no significant difference between missing and non-missing groups. In checking for outliers, Mahalanobis distance testing indicated just a single respondent with standardised residuals +/- three standard deviations from the predicted residual. Harman's one-factor test was conducted to test the presence of common method bias (Podsakoff *et al.* 2003). All scale variables were entered into an exploratory principal components factor analysis (PCA) and principal axis factoring (PAF) and subjected to an oblique rotation to identify how many factors are required to account for variance. Both PCA and PAF revealed the presence of 15 factors with eigenvalues >1.0 rather than a single factor. Of the 72.8% of variance explained by the 15 factors, only 25.6% was explained by the first factor, indicating no general factor is present (Aulakh and Gencturk, 2000; Podsakoff *et al.* 2003). These results suggest that the risk of common method bias is minimal. Data exhibit multivariate normality, with limited skew (-.705) and kurtosis (.448), whilst the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy (0.926) and Bartlett's Test of Sphericity (<.000) indicate the suitability of proceeding with factor analysis.

## Data Analysis

### *Gap-measure of internal service quality*

The 33 survey items used in the gap measure of internal service quality were subjected to exploratory factor analysis and extracted using principal axis factoring and oblique rotation. Total variance extracted is 70.9%, whilst common variance extracted is 64%. Based on the sample size of 274 in this research, all loadings greater than .35 are considered significant (Hair *et al.* 2006). Of the 33 gap items in the original factor solution, 30 were retained following purification for non-loading (*visual appeal*) and cross-loading (*talking user's language* and *encouraging feedback*). Table III shows the final factor solution for the internal service quality scale based on gap scores, with details of factor loadings, variance explained, and eigenvalues. Table IV shows the correlation matrix and descriptive statistics for the scale.

**Table III. Factor analysis of gap-measure of internal service quality**

Items	1	2	3	4	5	6
<b>1. Professionalism</b>						
support availability	.83					
support reliability	.78					
support responsiveness	.89					
support knowledge	.82					
support flexibility	.71					
problem resolution	.75					
Confidentiality	.82					
Friendliness	.86					
concern shown	.91					
<b>2. Processing</b>						
order processing speed		.66				
ease of authorisation		.54				
orders to supplier speed		.90				
order lead-time		.80				
processing complex orders		.49				
on-time delivery		.80				
order accuracy		.69				
system security		.56				
<b>3. Training</b>						
timely training			.88			
appropriate training			.98			

information provision	.65					
<b>4. Specification</b>						
FMS integration	.66					
invoice reconciliation	.64					
system configurability	.48					
reporting capability	.71					
<b>5. Content</b>						
loaded suppliers	.73					
loaded catalogues	.87					
ease of search	.47					
<b>6. Usability</b>						
system availability	.40					
screen loading speed	.73					
ease of navigation	.62					

<b>Variance explained</b>	40.79	12.08	5.26	4.87	4.34	3.58
<b>Eigenvalues</b>	12.24	3.62	1.58	1.46	1.30	1.07

**Table IV. Correlation matrix and descriptive statistics for gap-measure of internal service quality<sup>a, b</sup>**

Variable	Mean	S.D	1	2	3	4	5	6	7	8	9	10
<b>1. Professionalism</b>	-1.17	1.36	<b>.95</b>									
<b>2. Processing</b>	-1.22	1.07	.49	<b>.90</b>								
<b>3. Training</b>	-1.82	1.69	.63	.42	<b>.92</b>							
<b>4. Specification</b>	-1.69	1.35	.48	.64	.35	<b>.82</b>						
<b>5. Content</b>	-1.94	1.61	.37	.51	.30	.48	<b>.80</b>					
<b>6. Usability</b>	-1.63	1.26	.45	.60	.43	.56	.49	<b>.75</b>				
<b>7. OEPQ Rating</b>	4.81	1.26	.69	.58	.56	.54	.42	.41	-			
<b>8. IT skills</b>	5.15	1.11	.09	-.01	.16	-.05	-.02	-.04	.07	-		
<b>9. Experience</b>	7.45	4.85	.04	-.08	.02	-.11	-.04	-.04	.07	-.04	-	
<b>10. Size</b>	13968	13010	.01	-.02	-.04	.03	.08	.02	.02	.02	.09	-

<sup>a</sup> Correlation coefficient of .30 or greater are significant at  $p < 0.01$ ,  $n=255$ .

<sup>b</sup> Cronbach alpha shown in bold on diagonal

Given the fact that the research was not longitudinal (test-retest) and there is no alternative construct measure (parallel forms), assessment of reliability focuses on internal consistency (Flynn *et al.* 1990). Cronbach alphas for the six factors range from .75 to .95, and exceed the recommended cut-off points of .60 and .70 (Nunally, 1978). The overall alpha for the scale is .95. These results, combined with item-to-

total scores (.54 to .90, average .72), indicate a high level of internal consistency between items making up each factor.

The high reliabilities and clear factor structure provide support for trait validity of the gap-based measure of internal service quality. However, this is not sufficient in assessing the extent to which a scale captures the latent construct (Churchill, 1979). Content validity cannot be determined statistically, but rather by experts with reference to experience and literature (Sekaran, 2003). The items used to measure internal service quality (Brandon-Jones, 2008) draw on a wide range of service quality, internal service, information systems and e-service literature. The resulting scale appears to accurately reflect the construct, thus exhibiting good content validity.

Construct validity measures the extent to which a scale is a good operational definition of a construct and can be split into two elements. Convergent validity is established when variables load on a single factor and correlate with other variables in their assigned factors (Bagozzi, 1981). Discriminate validity is indicated if the factors and variables are truly different from one another (Carman, 1990). The rules of variable convergence and discrimination hold good for this data set. The factor analysis reveals that of the original 33 variables, 30 have high loadings on a single factor. In addition, the scale exhibits high alphas and high item-to-total scores. Finally, the high correlation between internal service factors (Table IV) provides additional evidence of construct validity (cf. Parasuraman *et al.* 1988).

Predictive validity is derived by examining the power of scale to predict scores on a separate criterion (Flynn *et al.* 1990). It is established when the measure differentiates individuals on a criterion as predicted (Sekaran, 2003). Predictive validity of the factors has been examined using multiple linear regression, with data controlled for e-procurement experience, IT skills, and organisational size (Table V).

The statistical power of the regression model is partly determined by the number of independent variables and the significance level chosen. For this research, using the six factors as independent variables and specifying a .01 significance level, the sample of 274 will detect  $R^2$  values of around 7% and greater. Assuming a representative sample, the ratio of observations to independent variables should always be greater than 5-to-1 and ideally 20-to-1 (Hair *et al.* 2006). In this research, the ratio of observations to independent variables is 45.7-to-1. Because our data are technically ordinal (i.e. 1-7 Likert scales), we ran an ordered logit model to ensure that both the significance pattern of coefficients and significance of factors was

identical to that produced by a multiple regression. This was the case and therefore the more commonly applied multiple regression approach is presented below.

**Table V. Results of regression analysis for gap-based measure of ISQ on OEPQ**

	<i>Model 1 – OEPQ rating</i>			
	Step 1		Step 2	
	$\beta$	t	$\beta$	t
<b>Controls</b>				
E-procurement experience	.07	1.06	.08	1.96
IT skills	.07	1.13	.02	.53
Organisational size	.01	.15	.01	.18
<b>Main effects</b>				
Professionalism			.40***	7.00
Processing			.23***	3.83
Training			.15**	2.64
Specification			.16**	2.77
Content			.06	1.18
Usability			-.05	.94
$\Delta R^2$	.01		.58***	
$\Delta F$	.79		57.45***	
Overall $R^2$	.01		.59	
Adjusted $R^2$	-.01		.57	
Overall model $F$	.80		38.93***	

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

The six factors explain 57% of variance in independent construct, the *overall e-procurement quality rating (OEPQ)*. The *professionalism* dimension was the most important predictor of *OEPQ ratings*. This dimension is concerned with the ongoing support provided to internal customers and emphasises support availability, responsiveness, reliability, and flexibility in solving problems. In addition, the attitude shown by support personnel is also considered. The dominance of *professionalism* is perhaps unsurprising given the large number of studies that emphasise the critical importance of providing adequate help to individuals who encounter problems with an internal service (cf. Bruhn, 2003; Cavinato, 1987; Chaston, 1994; Finn *et al.* 1996; Grönroos, 1988; Johnston and Silvestro, 1990; Kang *et al.* 2002; McDermott and Emerson, 1991; Parasuraman *et al.* 1985, 1988; Pitt *et al.* 1995, 1997; Rossler and Hirsz, 1996; Van Dyke *et al.* 1997; Young and Varble, 1997).

Whilst *content* and *usability* are correlated to the *OEPQ rating*, they produce only a marginal improvement to the regression model and are not statistically significant. This is because the predictive power of additional independent variables is not only determined by their correlation to the dependent variable, but also their correlation to

other independent variables in the model. As such, the value of *content* and *usability* factors is limited by their strong relationship with *professionalism*, *processing*, *training* and *specification*. However, it is important to avoid the conclusion that these factors are inconsequential in driving perceptions of internal service quality simply because they are not significant in this regression model.

In summary, the internal service measure based on gap scores appears to meet all the criteria to be considered reliable and valid. Our analysis now moves on to assess a scale based on perceptions-only data.

*Perceptions-only measure of internal service quality*

During data analysis of the perceptions-only measure of internal service quality, choices of method selection, factor design, retention of factors, extraction, rotation, interpretation, scale purification, creation of summated scales, and validation, were identical to those used in the gap-based measure assessment. Table VI shows the final factor solution for the perceptions-measure of internal service quality, with details of factor loadings, variance explained, and eigenvalues. Table VII shows the correlation matrix and descriptive statistics for the perceptions-based measure of internal service quality. Of the 33 perceptions items entered into the factor analysis, four were deleted during scale purification due to non-loading (*visual appeal*) and cross-loading (*talking user's language*, *encouraging feedback*, and *order accuracy*). The remaining 29 items load on a single factor. The scale explains 75.43% of total variance and 69.32% of shared variance.

**Table VI. Factor analysis of perceptions-measure of internal service quality**

Items	1	2	3	4	5	6
<b>1. P-Professionalism</b>						
support availability	.80					
support reliability	.88					
support responsiveness	.91					
support knowledge	.89					
support flexibility	.74					
problem resolution	.78					
Confidentiality	.85					
Friendliness	.92					
concern shown	.95					
<b>2. P-Processing</b>						

order processing speed	.73											
ease of authorisation	.56											
orders to supplier speed	.97											
order lead-time	.75											
processing complex orders	.42											
on-time delivery	.68											
system security	.53											
<b>3. P-Training</b>												
timely training	.90											
appropriate training	.98											
information provision	.67											
<b>4. P-Specification</b>												
FMS integration	.60											
invoice reconciliation	.68											
system configurability	.61											
reporting capability	.66											
<b>5. P-Content</b>												
loaded suppliers	.85											
loaded catalogues	.92											
Ease of search	.70											
<b>6. P-Usability</b>												
system availability	.54											
screen loading speed	.86											
ease of navigation	.49											

<b>Variance explained</b>	46.09	12.13	4.96	4.86	3.87	3.53
<b>Eigenvalues</b>	13.37	3.52	1.44	1.41	1.12	1.02

**Table VII. Correlation matrix and descriptive statistics for perceptions-measure of internal service quality <sup>a, b</sup>**

Variable	Mean	S.D	1	2	3	4	5	6	7	8	9	10
<b>1. P-Professionalism</b>	<b>5.25</b>	1.29	<b>.97</b>									
<b>2. P-Processing</b>	5.36	1.09	.59	<b>.90</b>								
<b>3. P-Training</b>	4.67	1.57	.66	.45	<b>.93</b>							
<b>4. P-Specification</b>	4.53	1.27	.52	.69	.40	<b>.85</b>						
<b>5. P-Content</b>	4.20	1.36	.42	.58	.33	.60	<b>.86</b>					
<b>6. P-Usability</b>	5.00	1.20	.46	.66	.43	.57	.53	<b>.78</b>				
<b>7. OEPQ Rating</b>	4.81	1.26	.78	.62	.62	.60	.52	.48	-			
<b>8. IT skills</b>	5.15	1.11	.05	-.03	.18	.01	.04	-.08	.07	-		
<b>9. Experience</b>	7.45	4.85	.06	-.06	.01	-.10	-.03	-.02	.07	-.04	-	
<b>10. Size</b>	13968	13010	-.01	-.03	-.05	.001	.01	.007	.02	.02	.09	-

<sup>a</sup> Correlation coefficient of .33 or greater are significant at  $p < 0.01$ ,  $n=255$ .

<sup>b</sup> Cronbach alpha shown in bold on diagonal

Internal reliability is indicated by Alpha coefficients which range from .78 to .97 for the six factors and .96 for the entire scale. These results, combined with item-to-total scores (.60 to .92), indicate a high level of internal consistency between items making up each factor and the scale as a whole. In terms of construct validity, the rules of variable convergence and discrimination (Bagozzi, 1981) hold good for the perceptions-only data. Of the 33 variables entered into the perceptions-only analysis, 29 have high loadings on a single factor. In addition, high scale alphas, item-to-total scores, and correlations between internal service factors provide evidence of construct validity. Predictive validity has been assessed using linear regression, with data controlled for e-procurement experience, IT skills, and organisational size (Table VIII).

**Table VIII. Results of regression analysis for perceptions-based measure of ISQ on OEPQ**

	<i>Model 1 – OEPQ rating</i>			
	Step 1		Step 2	
	$\beta$	t	$\beta$	t
<b>Controls</b>				
E-procurement experience	.07	1.06	.06	1.63
IT skills	.07	1.13	.02	.48
Organisational size	.01	.15	.03	.76
<b>Main effects</b>				
P-Professionalism			.49***	9.51
P-Processing			.14*	2.55
P-Training			.15**	3.03
P-Specification			.15**	2.70
P-Content			.11*	2.33
P-Usability			-.03	.70
$\Delta R^2$	.01		.68***	
$\Delta F$	.80		92.35***	
Overall $R^2$	.01		.70	
Adjusted $R^2$	-.002		.69	
Overall model $F$	.80		62.41***	

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

The six factors explain 69% of variance in the independent construct, the *OEPQ rating*. *Professionalism* is again the most important factor in the regression. Whilst *usability* is correlated to the *OEPQ rating*, it is not statistically significant. In summary, the internal service measure based on perceptions-only scores appears to meet all the criteria to be considered reliable and valid. Our analysis now moves to a comparison of the two measures of internal service quality.

*Comparing the two measures of internal service quality*

The various tests carried out indicate that both a gap- and a perceptions-only approach to measuring internal service quality produce scales with high levels of reliability and validity (Table IX).

**Table IX. Summary of scales – gap-measure versus perceptions-measure**

Gap-measure of internal service quality		Perceptions-measure of internal service quality
<b>Reliability</b>		
.75 - .95	Factor alpha range	.78 - .97
.95	Scale alpha	.96
.716	Item-to-total average	.758
High	<b>Content validity</b>	High
<b>Construct validity</b>		
30 of 33	Variables included in factor solution	29 of 33
.726	Average loading on assigned factor	.748
90.91%	Variables loading on single factor	87.9%
<b>Predictive validity</b>		
Adjusted R <sup>2</sup> .549	Regression: ISQ factors to OEPQ	Adjusted R <sup>2</sup> .665
165.83 from 402.97	Residuals (unaccounted variation)	122.39 from 402.97

Our analysis indicates that the perceptions-only scale of internal service quality outperforms the gap-based scale in a number of ways. In most areas, the improvement in performance is marginal, including factor and scale alphas, item-to-total scores, and factor loadings. However, considering predictive power, the perception-only scale of internal service quality explains 11.6% more variation in the dependent variable than the gap-based scale, which may be important in some contexts.

Our data highlight differences in the order of internal service factors and items depending on the use of a gap-based or perceptions-only approach. At a factor level, *professionalism* is considered the best area of service provision based on gap-scores (Table IV), whilst *processing* is considered the best using perceptions-scores (Table VII). Considering individual items (Appendix 1), the *confidentiality* item is ranked 4<sup>th</sup> on the basis of its gap-score, but 9<sup>th</sup> on the basis of its perceptions-score, whilst *visual appeal* is ranked 5<sup>th</sup> based on its gap-score, but 29<sup>th</sup> if the perceptions-only approach is applied. At the other end of the scale, *loaded catalogues* and *system configurability*

are the second and fourth worst performing internal service items when using perceptions-scores, but only the ninth and twelfth worst when gap-scores are applied.

## **Discussion**

On the basis of our analysis, we can now reflect on the research questions posed earlier.

### *How reliable and valid is a gap-based measure of internal service quality?*

This study confirmed the reliability and validity of the gap-based measure of internal service quality. One of the critiques of the gap-based measures noted in the literature review is that external customers tend to inflate expectation scores based on social norms. This study identified generally high expectation levels, suggesting that, in the same way as external customers, internal customers may be prone to expectation inflation. However, our findings contrast with Large and König (2009) who report lower and more varied internal service expectations compared with many external service studies. This suggests that the problem of inflated expectations scores occurs in some internal customer services but not in others. Therefore, we conclude that generalisations regarding the danger of expectation inflations cannot yet be made for internal service contexts.

It is also argued that in external customer contexts gap-based measures have higher diagnostic value than perceptions-only measures, and that difference scores can better pinpoint areas of deficiency within an organisation (Parasuraman *et al.*, 1994b, Pitt *et al.* 1995, 1997; Dean, 1999). This study suggests that this is also true in internal services: our data reveal significant differences in the order of internal service items and factors based on gap-scores as opposed to perceptions-scores. In these cases, the gap-scores are likely to be the more useful measure in terms of identifying improvement priorities, since they enable managers to target improvements of those aspects of service where internal customer expectations are high.

The respondents in this study had never before participated in a survey of their views on internal service quality. A priority for management therefore was to gain some understanding of these customers' expectations as well as their perceptions, and clearly the gap-based measure provided richer information in this respect. Moreover regular monitoring using gap-based measurement would facilitate an understanding

of changes in both expectations and perceptions over time. This is particularly important in industries where expectations are poorly understood – a common problem in internal services where, as was argued earlier, there has been a dearth of research into internal customers.

In external customer contexts it is recognised that the gap-approach is particularly pertinent in turbulent competitive arenas where customer expectations are highly dynamic and constantly changing in response to new competitive offerings. In this respect, internal customer relationships might generally be expected to be more stable than external customer relationships and therefore there may be less of an imperative to use gap-based measures in internal services. Indeed this was true of the internal services in this study: the working environment was stable, staff turnover was low and many of the employees had worked there for many years. However, the turbulence of the internal service market must be judged by the managers who are implementing the measurement system. In organisations where there has been significant organisational change, high staff turnover and general disruption to service activities and processes, it may well be necessary to measure changes in internal customer expectations as well as their perceptions.

*How reliable and valid is a perceptions-only measure of internal service quality?*

The perceptions-only measure marginally outperforms the gap-based measures in terms of reliability and validity. When the focus of study is on prediction of related constructs, the perceptions-only approach appears to be particularly strong. Furthermore, the practical advantages of the substantially reduced questionnaire length, compared to the gap approach, should be recognised. In this study, survey weariness of staff was not a problem: the respondents cooperated well with the survey process, in fact many were pleased to be given the opportunity to express their views on e-procurement service provision. However, in internal services where there is more reluctance to engage, for example, due to previous participation in surveys or a perception that feedback does not result in improvement actions, questionnaire length will be a more significant issue and the shorter perceptions-only questionnaire may be advantageous. Indeed in organisations where expectations are considered to be relatively stable, it may only be necessary to measure expectations separately once every three years, as Carman (1990) advises with regard to external quality. This could reduce the likelihood of boredom setting in during questionnaire completion,

thus improving response rates and heightening confidence in subsequent data analysis (Babakus and Boller 1992).

However, this study has also highlighted some of the drawbacks of the perceptions-only measure; in particular, the risk that adopting a perceptions-only measure of internal service quality can result in the misdiagnosis of improvement priorities. For example, using perceptions data from this study, an operations manager would have likely focused improvement efforts on *visual appeal*, *loaded catalogues*, and *loaded suppliers*, despite the fact that gap scores indicate that *timely training*, *appropriate training*, and *ease of navigation* all have significantly larger gaps between what is expected and what is being delivered (See appendix 1). Indeed the internal customers' expectations of the latter items were consistently higher than their expectations of *visual appeal*, *loaded catalogues* and *loaded suppliers*. Thus prioritising improvements based on the gap measure might have more of an impact on internal customers' perceptions of quality than prioritisation based on the perceptions-only measure.

Furthermore, if internal service quality is measured longitudinally and performance trends are to be analysed, the perceptions-only measure is limited in that, unlike the gap-based measure, it does not enable managers to interpret sudden or unexpected changes in internal quality. For example, if perceived internal quality is in decline, the perceptions-only approach fails to reveal whether this is due to reduced internal service levels or a rise in expectations.

#### *What are the benefits and limitations of each approach?*

There is clearly a trade-off between the data richness and diagnostic value of the paired-statement gap approach, compared with the marginally higher reliability, validity and significant collection efficiencies gained from the single-statement perceptions-only approach. If the development of a shared understanding of internal customer expectations is a managerial priority, then the gap approach will provide data which can be used to help internal suppliers understand how customer expectations vary between different aspects of service and change over time. The more turbulent the organisational environment and the less stable internal customer expectations are, the greater the need to separately measure expectations and quantify the expectation-perception gap. However, the pragmatic advantages of the perceptions-only approach, with substantially reduced questionnaire length, are

significant and should not be under estimated in internal service environments. The attitude of internal customers towards the measurement process can have a major impact on their commitment to the process, their willingness to engage and their generosity in making time for the questionnaire completion. This will in part depend upon whether similar surveys have taken place in the past, and the perceived outcomes of these endeavours.

On the basis of the above discussion the benefits and limitations of the two approaches are summarised in Table X. Ultimately, the choice of method must depend on the managerial purpose of the internal quality measurement system. This moves the debate on measuring internal service quality away from evaluations of the effectiveness of gap-based versus perceptions-only approaches, and towards consideration of the operational contexts in which each approach might be appropriate.

**Table X: Evaluation of gap-measure versus perceptions-measure of internal service quality**

	<b>Gap-measure of ISQ</b>	<b>Perceptions- measure of ISQ</b>
<b>BENEFITS</b>	<ul style="list-style-type: none"> <li>• Valid and reliable</li> <li>• Data richness</li> <li>• Improved understanding of expectations</li> <li>• Increased diagnostic value: effective in identifying improvement priorities</li> </ul>	<ul style="list-style-type: none"> <li>• Marginally increased reliability and validity</li> <li>• Increased predictive power</li> <li>• Higher response rates</li> </ul>
<b>LIMITATIONS</b>	<ul style="list-style-type: none"> <li>• Lengthy questionnaires</li> <li>• Respondent boredom</li> <li>• Lower response rate</li> <li>• Data proliferation</li> </ul>	<ul style="list-style-type: none"> <li>• Failure to monitor changes in expectations</li> <li>• Over-inflation or upward-bias of customer service ratings</li> <li>• Difficulty of interpreting unexpected changes in perceived quality</li> <li>• Potential misdiagnosis of improvement criteria</li> </ul>

*Limitations and suggestions for future research*

This research is limited by the fact that it was carried out in one particular type of internal service – the provision of e-procurement software, training and user support. In line with other scale development and assessment studies (cf. Parasuraman *et al.* 1988; Reynoso and Moores, 1995; Finn *et al.* 1996), data were collected from a small number of organisations. It was not deemed appropriate to survey internal customers in a broader range of settings until the proposed measures of internal service quality

had been validated in the original research setting. However, there is clearly a need to further test alternative measurement approaches in different internal service contexts in order to refine our understanding of internal service quality measurement. Testing would also benefit from examining predictive validity of alternative measures against other dependent constructs, such as loyalty or complaints, as well as overall satisfaction (in this study the ‘overall e-procurement quality rating’).

## **Conclusion**

To conclude, this study has compared the gap-based and perceptions-only measures of internal service quality. Both approaches can be justified theoretically, and testing has established that both can be operationalised in ways which are reliable and valid. The study, combined with the contributions from the literature, has generated some understanding of the specific conditions in which the two approaches might be appropriate. The debate as to which approach is better is therefore superseded by what is perhaps a more productive perspective: one which aims to develop a better understanding of the factors that influence appropriate selection of internal service quality measures.

## **References**

- Ahmed, P.K. and Rafiq, M. (2000), “Advances in the internal marketing concept: definition, synthesis and extension”, *Journal of Services Marketing*, Vol. 14, No. 6, pp. 449-462
- Albrecht, K. and Bradford, L. J. (1990), *The Service Advantage: How to Identify and Fulfil Customer Needs*, Richard D. Irwin, Homewood, IL.
- Aulakh, P. S. and Gencturk, E. F. (2000), “International principal-agent relationships: control, governance and performance”, *Industrial Marketing Management*, Vol. 29, No. 6, pp. 521–538.
- Auty, S. and Long, G. (1999), “‘Tribal Warfare’ and gaps affecting internal service quality”, *International Journal of Service Industry Management*, Vol. 10, No. 1, pp. 7-18.
- Babakus, E. and Boller, G. W. (1992), “An Empirical Assessment of the SERVQUAL Scale”, *Journal of Business Research*, Vol. 24, No. 2, pp. 253-268.
- Bagozzi, R. O. (1981), “Attitudes, intentions, and behaviour: a test of some key hypothesis”, *Journal of Personality and Social Psychology*, Vol. 41, pp. 607-627.
- Boshoff, C. and Mels, G. (1995), “A causal model to evaluate the relationships among supervision, role stress, organizational commitment and internal service quality”, *European Journal of Marketing*, Vol. 29, No. 2, pp. 23-42.

- Bouman, M. and van der Wiele, T. (1992), "Measuring Service Quality in the Car Service Industry: Building and Testing an Instrument", *International Journal of Service Industry Management*, Vol. 3, No. 4, pp. 4-16.
- Brady, M. K., Cronin, J. J. and Brand, R. R. (2002), "Performance-only measurement of service quality: A replication and extension", *Journal of Business Research*, Vol. 55, No. 1, pp. 17-31.
- Brandon-Jones, A. (2006), *E-procurement Quality: Exploring and measuring the construct at a tactical level in the public sector*, PhD thesis, University of Warwick, UK.
- Brandon-Jones, A. (2008), The EPQ scale: a multi-item measure of perceived e-procurement quality, *Proceedings of the 19th annual conference of the Production and Operations Management Society*, La Jolla, California
- Brandon-Jones, A., Ramsey, J., Wagner, B. (2010), "Trading Interactions: Supplier Empathy, Consensus and Bias", *International Journal of Operations and Production Management*, forthcoming.
- Brooks, R.F., Lings, I. and Botschen, M. (1999), "Internal Marketing and Customer Driven Wavefronts", *The Service Industries Journal*, October, pp. 49-67.
- Bruhn, M. (2003), "Internal Service Barometers," *European Journal of Marketing*, Vol. 37, No. 9, pp. 1189-1204.
- Buttle, F. (1996), "SERVQUAL: Review, critique, research agenda", *European Journal of Marketing*, Vol. 30, No. 1, pp. 8-33.
- Carman, J. M. (1990), "Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions", *Journal of Retailing*, Vol. 66, pp. 33-55.
- Cavinato, J. (1987), "Purchasing Performance: What Makes The Magic?", *Journal of Purchasing and Materials Management*, Vol. 23, No. 3, pp. 10-16.
- Chaston, I. (1994), "Internal customer management and service gaps within the UK manufacturing sector", *International Journal of Operations and Production Management*, Vol. 14, No. 9, pp. 45-57.
- Churchill, G. A. Jr. (1979), "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, Vol. 16, February, pp. 64-73.
- Clow, K. E. and Vorhies, D. W. (1993), "Building a competitive advantage for service firms: measurement of consumer expectations of service quality, *Journal of Services Marketing*, Vol. 7, Issue 1, pp. 22-33.
- Cronin, J. and Taylor, S. (1994), "SERVPERF versus SERVQUAL: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality", *Journal of Marketing*, Vol. 58, January, pp. 125-131.
- Cronin, J. J. and Taylor, S. (1992), "Measuring Service Quality: A Reexamination and Extension", *Journal of Marketing*, Vol. 56, July, pp. 55-68.

- Croom, S. and Brandon-Jones, A. (2007), "Progress on E-Procurement: Experiences from Implementation in the UK Public Sector", *Journal of Purchasing and Supply Management*, Vol. 13, No. 4, pp. 294-303
- Davis, T. R. V. (1991), "Internal service operations: strategies for increasing their effectiveness and controlling their cost", *Organizational Dynamics*, Vol. 20, Autumn, pp. 5-22.
- Dean, A. (1999), "The applicability of SERVQUAL in different health care environments", *Health Marketing Quarterly*, Vol. 16, No. 3, pp. 1-21.
- Deming, W.E. (1986), *Out of the Crisis*, Centre for Advanced Engineering Study, Massachusetts Institute of Technology, Cambridge.
- Easterby-Smith, M., Thorpe, R. and Lowe, A. (1997), *Management Research: An Introduction*, Sage Publications, London.
- Eisenhardt, K.M. (1989), Building theories from case study research, *Academy of Management Review*, 14, 4, October, pp. 532-550
- Farner, S., Luthans, F. and Sommer, S. (2001), "An empirical assessment of internal customer service", *Managing Service Quality*, Vol. 11, No. 5, pp. 350-358.
- Finn, D., Baker, J., Marshall, G. and Anderson, R. (1996), "Total Quality Management and Internal Customers: Measuring Internal Service Quality", *Journal of Marketing Theory and Practice*, Vol. 4, No. 3, pp. 36-51.
- Flynn, B.B., Sakakibara, S., Schroeder, R.G., Bates, K.A. and Flynn, J.E. (1990), "Empirical Research Methods in Operations Management", *Journal of Operations Management*, Vol. 9, No. 2, pp. 250-284.
- Frost, F. and Kumar, M. (2000), "INTSERVQUAL - an internal adaptation of the GAP model in a large service organisation", *Journal of Services Marketing*, Vol. 14, No. 5, pp. 358-377.
- George, W.R. (1990), "Internal marketing and organizational behavior: a partnership in developing customer-conscious employees at every level", *Journal of Business Research*, Vol. 20, No. 1, pp. 63-70.
- Gilbert, G.R. (2000), "Measuring internal customer satisfaction", *Managing Service Quality*, Vol. 10, No. 3, pp. 178-185.
- Gremler, D.D., Bitner, M.J. and Evans, K.R. (1994), "The Internal Service Encounter", *International Journal of Service Industry Management*, Vol. 5, No. 2, pp. 34-56.
- Grönroos, C. (1988), "Service Quality: the Six Criteria of Good Perceived Quality", *Review of Business*, Vol. 9, No. 3, pp. 10-13.
- Hair, J.F. Jr., Anderson, R.E., Tatham, R.L. and Black, W.C. (2006), *Multivariate Analysis*, 6th edition, Prentice-Hall, London
- Hallowell, R., Schlesinger, L. and Zornitsky, J. (1996), "Internal Service Quality, Customer and Job Satisfaction: Linkages and Implications for Management", *Human Resource Planning*, Vol. 19, No. 2, pp. 20-31.

- Hendrick, T. and Ruch, W. (1988), "Determining Performance Appraisal Criteria for Buyers", *Journal of Purchasing and Materials Management*, Vol. 24, No. 2, pp. 18-26.
- Heskett, J.L., Jones, T.O., Loveman, G.W., Sasser Jr., W.E. and Schlesinger, L.A. (1994), "Putting the service profit chain to work", *Harvard Business Review*, March - April, pp. 164- 174
- Heskett, J. L., Sasser, W. E. and Schlesinger, L. A. (1997), *The Service Profit Chain*, Free Press, New York.
- Hill, F. and McCrory, M. (1997), "An attempt to measure service quality at a Belfast maternity hospital: Some methodological issues and some results", *Total Quality Management*, Vol. 8, No. 5, pp. 229-242.
- Howard, M., Lewis, M., Miemczyk, J. and Brandon-Jones, A. (2007), "Implementing supply practice at Bridgend Engine Plant. The influence of institutional and strategic choice perspectives", *International Journal of Operations and Production Management*, Vol. 27, No. 7, pp. 754-776.
- Iacobucci, D., Grayson, K.A. and Omstrom, A.L. (1994), "The calculus of service quality and customer satisfaction: theoretical and empirical differentiation and integration", in Swartz, T.A., Bowen, D.E., Brown, S.W. (Eds.), *Advances in Services Marketing and Management*, JAI Press, Greenwich, CT, Vol. 3, p. 1-68.
- Ishikawa, K. (1985), *What is Total Quality? The Japanese Way*, Prentice Hall, Englewood Cliffs
- Johnston, R. (1999), "Service operations management: return to roots", *International Journal of Operations and Production Management*, Vol. 19, No. 2, pp. 104-124.
- Johnston, R. (2005), "Service operations management: from the roots up", *International Journal of Operations and Production Management*, Vol. 25, No. 12, pp. 1298-1309.
- Johnston, R. Silvestro R. (1990), "The determinants of service quality - a customer-based approach", in *The Proceedings of the Decision Science Institute Conference*. San Diego, CA.
- Juran, J.M. (1989), *Juran on Leadership for Quality: an Executive Handbook*, The Free Press, New York.
- Kang, G-D., James, J. and Alexandris, K. (2002), "Measurement of internal service quality: application if the SERVQUAL battery to internal service quality", *Managing Service Quality*, Vol. 12, No. 5, pp. 278-291.
- Koska, M.T. (1992), "Surveying customer Needs, Not Satisfaction, Is Crucial to CQI", *Hospitals*, Vol. 66, No. 21, pp. 50-53.
- Kuei, C-H. (1999), "Internal Service Quality - an empirical assessment", *The International Journal of Quality and Reliability Management*, Vol. 16, No. 8, pp. 783-788.
- Large, R.O. and König, T. (2009), "A gap model of purchasing's internal service quality: concept, case study and internal survey", *Journal of Purchasing and Supply Management*, Vol. 15, pp. 24-32
- Lewis, B.R. and Gabrielson, G.O.S. (1995), "An Intra-organisational Approach towards the Implementation of Service Quality Management", *Manchester School of Management*.

- Marshall, G., Baker, J. and Finn, D. (1998), "Exploring internal customer service quality", *Journal of Business and Industrial Marketing*, Vol. 13, No. 4/5, pp. 381-392.
- McDermott, L. and Emerson, M. (1991), "Quality and Service for Internal Customers", *Training and Development Journal*, Vol. 45, No. 1, pp. 61-64.
- Nunnally, J.C. (1978), *Psychometric Theory*, 2nd edition, McGraw-Hill, New York.
- Paraskevas, A. (2001), "Internal service encounters in hotels: An empirical study", *International Journal of Contemporary Hospitality Management*, Vol. 13, No. 6, pp. 285-293.
- Parasuraman, A., Zeithaml, V. and Berry, L. (1985), "A conceptual model of service quality and its implications for future research", *Journal of Services Marketing*, Vol. 49, Fall, pp. 41-50.
- Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988), "SERVQUAL: A Multiple-Item Scale for measuring consumer perceptions of service quality", *Journal of Retailing*, Vol. 64, Spring, pp. 12-40.
- Parasuraman, A., Zeithaml, V.A., and Berry, L.L. (1994a), "Reassessment of Expectation as a comparison standard in measuring service quality: Implication for further research", *Journal of Marketing*, Vol. 58, January, pp. 111-124.
- Parasuraman, A., Zeithaml, V. and Berry L. (1994b), "Alternative Scales for measuring service quality: A Comparative Assessment Based on Psychometric and Diagnostic Criteria", *Journal of Retailing*, Vol. 70, No. 3, pp. 201-230.
- Peterson, R.A. and Wilson, W.R. (1992), "Measuring Customer Satisfaction: Fact and Artifact", *The Journal of the Academy of Marketing Science*, Vol. 20, No. 1, pp. 61-71.
- Phillips, D. and Clancy, K. (1972), "Some effects of "Social Desirability" in Survey Studies", *American Journal of Sociology*, Vol. 77, pp. 921-940.
- Pitt, L.F., Kavan, R.T. and Bruce, C. (1995), "Service quality: A measure of information systems effectiveness", *MIS Quarterly*, Vol. 19, No. 2, pp. 173-188.
- Pitt, L.F., Watson, R.T. and Kavan, C.B. (1997), "Measuring information systems service quality: concerns for a complete canvas", *MIS Quarterly*, Vol. 21, No. 2, pp. 209-222.
- Podsakoff, P., MacKenzie, S., Lee, J. and Podsakoff, N. (2003), "Common method biases in behavioural research: a critical review of the literature and recommended remedies", *Journal of Applied Psychology*, Vol. 88, pp. 879-903.
- Ratcliffe-Smith, J. and Brooks, R. (1993), "Service from within", *The TQM Magazine*, Vol. 5, pp. 41-43.
- Reynoso, J. and Moores, B. (1995), "Towards the measurement of internal service quality", *International Journal of Service Industry Management*, Vol. 6, No. 3, pp. 64-83.
- Rossler, P.E. and Hirsz, A.B. (1996), "Purchasing's interaction with customers: the effects on customer satisfaction – a case study", *International Journal of Purchasing and Materials Management*, Vol. 32, No. 1, pp. 37-43.

- Sackett, P. and Larson, J., (1990), "Research strategies and tactics in industrial and organizational psychology", In Dunnette, M. and Hough, L., (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 419–489). Palo Alto, CA: Consulting Psychologists Press
- Scarpello, V. and Campbell, J. (1983), "Job satisfaction: Are all the parts there?", *Personnel Psychology*, Vol. 36, pp. 577–600.
- Sekaran, U. (2003), *Research Methods for Business: A Skill Building Approach*, 4th edition, Wiley and Sons, Inc., Chichester, UK
- Silvestro, R. (2005), "Applying gap analysis in the health service to inform the service improvement agenda", *The International Journal of Quality and Reliability Management*, Vol. 22, No. 3, pp. 215-233.
- Smith, A.M. (1995), "Measuring service quality: is SERVQUAL now redundant?", *Journal of Marketing Management*, Vol. 11, No. 2, pp. 257-276.
- Stanley, L.L. and Wisner, J.D. (2001), "Service Quality along the supply chain: implications for purchasing", *Journal of Operations Management*, Vol. 19, No. 3, pp. 287-306.
- Stauss, B. (1995), "Internal services: Classification and quality management", *International Journal of Service Industry Management*, Vol. 6, No. 2, pp. 62-79.
- Teas, K.R. (1993), "Expectations, performance evaluation, and consumers' perceptions of quality", *Journal of Marketing*, Vol. 57, October, pp. 18-34.
- Teas, K.R. (1994), "Expectations as a Comparison Standard in Measuring Service Quality: An Assessment of a Reassessment", *Journal of Marketing*, Vol. 58, January, pp. 132-139.
- Van Dyke, T.P., Kappelman, L.A., and Prybutok, V.R. (1997), "Measuring information systems service quality: Concerns on the use of the SERVQUAL questionnaire", *MIS Quarterly*, Vol. 21, June, pp. 195-209.
- Wanous, J., Reichers, A. and Hurdy, M. (1997), "Overall job satisfaction: How good are single-item measures?", *Journal of Applied Psychology*, Vol. 82 No. 2, pp. 247-252
- Young, J.A. and Varble, D.L. (1997), "Purchasing's performance as seen by its internal customers: a study in a service organisation", *International Journal of Purchasing and Materials Management*, Vol. 33, No. 3, pp. 36-41.

**Appendix 1. Internal service quality items, definitions, descriptive data, and rankings**

Variable	Definition	Expectations mean	Expectations standard deviation	Perceptions mean	Perceptions standard deviation	Gap mean	Gap standard deviation	Gap ranking (best to worst)	Perceptions ranking (1=best)
concern shown	extent to which support personnel are willing to listen and empathize	6.13	.971	5.46	1.32	-.67	1.49	1	4
friendliness	level of friendliness shown by support personnel in dealings with users	6.38	.80	5.70	1.24	-.69	1.35	2	2
system security	system ability to minimize risk of fraud or loss of financial information	6.73	.56	5.79	1.10	-.94	1.10	3	1
confidentiality	confidence that dealing with support personnel will be treated with discretion	6.35	.88	5.31	1.35	-1.04	1.59	4	9
visual appeal <sup>a+b</sup>	the visual appeal of the software	5.39	1.22	4.34	1.68	-1.05	1.78	5	29
system availability	ease of accessing the system, incorporating system and server reliability	6.60	.71	5.44	1.28	-1.15	1.43	6	5
orders to supplier speed	speed and reliability of getting orders to suppliers from using the system	6.72	.51	5.54	1.28	-1.17	1.30	7	3
support knowledge	technical competence of support personnel to deal with queries	6.54	.67	5.36	1.45	-1.18	1.55	8	7
order accuracy <sup>b</sup>	impact on level of accuracy from using the system	6.54	.91	5.30	1.42	-1.24	1.43	9	10
on-time delivery	impact on number of on-time deliveries from using the system	6.46	.92	5.19	1.39	-1.27	1.42	10	13
support responsiveness	speed of response to user queries	6.53	.66	5.23	1.56	-1.30	1.66	11	11
support flexibility	willingness to meet various demands of users	6.30	.85	5.00	1.52	-1.30	1.64	12	16
ease of authorisation	ease and speed of authorizing order requisitions from using the system	6.69	.51	5.39	1.48	-1.30	1.49	13	6
processing complex orders	system ability to process complex orders where requisitions and invoices often do not match	6.29	.91	4.96	1.43	-1.32	1.60	14	18
order lead-time	impact on time taken to deliver an order from using the system	6.42	.78	5.09	1.51	-1.33	1.58	15	15
order processing speed	impact on order processing speed from using the system	6.69	.54	5.33	1.34	-1.36	1.33	16	8
problem resolution	ability of support personnel to resolve problems	6.55	.67	5.18	1.31	-1.38	1.45	17	14
support availability	availability of support to deal with problems when users encounter difficulties	6.33	.90	4.93	1.65	-1.39	1.85	18	19
system configurability	extent to which workflow, budget links, authorization levels, reporting, and screen appearance can be customised	5.88	1.2	4.46	1.52	-1.42	1.66	19	27
support reliability	reliability of support personnel to get back to users when they say they will	6.62	.64	5.19	1.52	-1.43	1.60	20	12
reporting capability	variety of report options, ease of searching for management information, and ability to customize reports	6.09	1.05	4.61	1.37	-1.48	1.52	21	25
loaded catalogues	extent to which content is loaded on the system	5.69	1.3	4.21	1.45	-1.48	2.01	22	32
information provision	provision of up-to-date information about system updates, new catalogues, suppliers, procurement rules, user tips	6.27	.86	4.71	1.56	-1.55	1.71	23	22
talking users' language <sup>a+b</sup>	communicating in a way that is easy to understand for users	6.53	.70	4.97	1.56	-1.56	1.72	24	17
screen loading speed	speed at which pages on the system load	6.54	.71	4.87	1.43	-1.66	1.56	25	20
encouraging	communicating in a way that is easy to	6.16	.91	4.37	1.67	-1.79	1.88	26	28

feedback <sup>a+b</sup>	understand for users								
loaded suppliers	extent to which suppliers are loaded on the system	6.15	1.09	4.32	1.52	-1.83	1.93	27	30
invoice reconciliation	system ability to 3-way match requisitions, orders, and invoices	6.61	.639	4.74	1.64	-1.87	1.70	28	21
timely training	provision of timely training by support personnel to users	6.53	.77	4.63	1.71	-1.90	1.90	29	24
appropriate training	provision of appropriate training by support personnel to users	6.65	.62	4.68	1.70	-1.97	1.81	30	23
FMS integration	system ability to work alongside legacy finance systems	6.29	1.05	4.23	1.56	-2.06	1.78	31	31
ease of navigation	ease with which users are able to find their way around the system	6.69	.53	4.60	1.53	-2.09	1.60	32	26
ease of search	ease of searching for suppliers and catalogues on the system	6.56	.76	3.99	1.63	-2.57	1.84	33	33

<sup>a</sup> item deleted during gap-based factor analysis

<sup>b</sup> item deleted during perceptions-based factor analysis

## About the authors



**Dr Alistair Brandon-Jones** is a Lecturer in Operations and Supply Management at the University of Bath and a visiting lecturer at Warwick Medical School. His main area of research focuses on supply strategy and for this work he is the UK lead member for the International Purchasing Survey ([www.ipsurvey.org](http://www.ipsurvey.org)) which explores the procurement processes and performance across the globe, in collaboration with a number of universities in Europe and the US. Another research interest is customer-centric service design. This work focuses on the important role which customers – either internal or external – can have in improving service delivery. Alistair is published in the *International Journal of Operations and Production Management*, *Journal of Purchasing and Supply Management*, and *Journal of Public Procurement*, and has a book, co-authored with Professor Nigel Slack, *Quantitative Analysis in Operations Management*, published by Pearson.



**Dr Rhian Silvestro** is an Associate Professor in Operations Management at the University of Warwick. Rhian has conducted service management research in a number of large, leading edge organisations including retail companies, banks, transport companies and call centres. She has acted as a consultant to ward managers in NHS hospital trusts, as well as NHS Direct, in the area of nurse scheduling and the computerisation of rostering systems. Rhian's work is published in journals which include *OMEGA International Journal of Management Science*, *International Journal of Operations and Production Management*, *International Journal of Service Industry Management*, *International Journal of Quality and Reliability Management*, *Design Management Journal*, *Health Services Management Research*, and *Journal of Advanced Nursing*. She is co-author of *Performance Measurement in Service Businesses*, published by CIMA.