

**008-0187**

# **The Role of Expectations in Measuring Internal Service Quality: A Contingency Perspective**

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

*Assistant Professor in Operations, Supply & Decision Analysis, Bath School of Management, Bath University, 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)*

*POMS 19th Annual Conference, La Jolla, California, U.S.A.*

*May 9 to May 12, 2008*

## **Abstract**

This paper examines the role of expectations in the measurement of perceived quality in internal service contexts. Our research examines the debates in the literature regarding both the theoretical and practical efficacy of expectations data in the measurement of internal service quality. Based on data collected from the internal customers of e-procurement systems in four organizations, two generic approaches to internal service quality measurement are tested empirically: a gap-based and a perceptions-only

---

<sup>1</sup> Corresponding Author

methodology. A third methodology, based on a single-statement gap measure is proposed as a further alternative approach which may be appropriate in some internal services contexts. The empirical testing, combined with contributions from the literature, generates some understanding of the specific organisational and managerial conditions in which the three approaches might be appropriate. The debate as to which methodology is better is therefore superseded by a perhaps more productive approach: one which aims to develop a better understanding of the contingencies which determine appropriate implementation.

**Keywords:** Internal service quality, contingency theory, disconfirmation theory, EPQ scale.

## **1. Introduction**

The debate as to how service quality should be modelled and operationalised into effective quality measurement systems has raged for some three decades. The discussion has centred around whether measurement should be based on disconfirmation theory, and operationalised with a measure of the gap between expectations and perceptions, or whether perception-only measures might be more reliable and efficacious. This debate has pertained mainly to external customer perceived quality.

There is a more limited but growing literature on internal service quality, and a move to evaluate the transferability of these approaches to internal service contexts. This gives rise to the same debate and there is a need to evaluate and compare gap-based and perception-only measures of perceived quality in this context. This paper reports the findings of an empirical study which tests the two approaches both in terms of their theoretical underpinnings and also in the light practical considerations regarding the design of measurement systems. Case studies and surveys were conducted in four UK organisations, based on the provision of e-procurement software systems to internal users.

## **2. Literature Review**

The disconfirmation paradigm, developed by such exponents as Oliver, (1980), Grönroos (1982), Smith and Houston (1982), Lewis and Booms (1983) and Cadotte *et al.* (1987), underpins the service quality literature of the past three decades. Disconfirmation theory states that customer satisfaction is determined by a comparison between previously held expectations and perceptions of performance. Disconfirmation is positive when

performance exceeds expectations and negative when it falls short of expectations; and the direction of disconfirmation determines whether a customer will be satisfied or dissatisfied with the service they receive (Ganesh *et al.* 2000). The disconfirmation model is now the dominant paradigm for the measurement of service quality (Brooks *et al.* 1999; Silvestro 2005).

Oliver (1999) argues that *assimilation* and *contrast* effects significantly influence the disconfirmation process. Assimilation effects relate to the importance of previously held expectations in anchoring performance assessments. Therefore, the extent and nature of customer experience is critical in the disconfirmation process (Cadotte *et al.* 1987). Additionally, assimilation effects are particularly important when the level of performance is ambiguous (Ganesh *et al.* 2000). Contrast effects relate to the extent to which customers magnify perceptions ratings in line with the direction of disconfirmation. Customers may over-emphasise positive or negative disconfirmation by indicating extreme performance ratings. High levels of customer involvement in the service process tend to increase contrast effects.

### ***Theoretical Concerns with Disconfirmation Theory***

Despite its prevalence, the disconfirmation model of service quality is not without its critics. First, there are concerns in defining a construct as the difference between two other constructs, and a number of authors have criticised the gap approach on the basis of its theoretical underpinnings (Carman, 1990, Babakus and Boller, 1992; Cronin and Taylor, 1992, 1994; Teas 1993a, 1993b, 1994; Brown *et al.* 1993; Peter *et al.* 1993, Brady *et al.* 2002;). They argue that a perceptions-only (i.e. direct / non-difference) approach based on attitudinal theory is more appropriate for measuring service quality. For example, Cronin and Taylor (1992, 1994) challenge the validity of the P-E approach, instead positing an unweighted performance-based measure of service quality – a view that is supported by Smith (1995).

Second, those applying disconfirmation theory to service quality measurement are sometimes accused of ‘muddying the waters’ in terms of the difference between customer satisfaction and service quality. Disconfirmation may be appropriate when measuring the transaction-specific concept of customer satisfaction. However, given that, ‘service quality is an overall evaluation similar to an attitude’ (Parasuraman *et al.* 1988, p15), there are questions regarding the theory’s applicability to service quality measurement. As such, disconfirmation theory is at odds with the ‘attitudinal’ model of service quality.

Third, Buttle (1996) notes the failure to assess customer expectations based on absolute rather than relative standards, means that service quality is assumed if there is no gap between expectations and perceptions. The fact that lower expectations make delivery of 'satisfactory service' easier is described by Grönroos as a 'service paradox'. Van Dyke *et al.* (1997) state that the use of difference scores in service quality measurement was originally an operational decision, rather than being theoretically based. They argue that, 'a direct measure of one's perceptions of service quality that is the outcome of this cognitive evaluation process seems more likely to yield a valid and reliable outcome'. Responding to such criticisms, Parasuraman *et al.* have strongly defended their approach to service quality measurement but ultimately recognise that the debate regarding the best way to define service quality remains inconclusive (Parasuraman *et al.* 1993, 1994a, 1994b).

### ***Practical Concerns with Disconfirmation Theory***

Whilst disconfirmation theory has been challenged on the basis of its conceptual underpinnings, concerns have also been raised about the *practical* merit of gap scores. A number of authors have questioned the additional value of gap scores (P-E) over perceptions-only data in predicting related dependent variables (overall quality, customer retention, recommendation etc). The question is, 'what additional information does the difference score provide over and above the perceptions scores?'

Iacobucci *et al.* (1994) support the disconfirmation perspective of service quality at a theoretical level, but argue that in practice gap scores may be very unreliable even if they are derived from reliable measures of expectations and perceptions. 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.' Cronin and Taylor (1994) state that, 'consumer perceptions, not calculations, govern behaviour'. Van Dyke *et al.* (1997) question whether gaps scores are useful in measuring the outcome of cognitive discrepancy.

Babakus and Boller (1992) state that correlations between SERVQUAL, overall ratings and complaint resolution are simply weaker versions of the correlations between perceptions and these dependent variables. A study by Parasuraman *et al.* (1993) compares the predictive power of difference scores and perceptions-only scores through multiple regression analysis. R<sup>2</sup> scores for SERVQUAL range from .51 to .71, whilst perceptions-

only scores are .72 to .81. Brown *et al.* (1993) also identify higher correlations between behavioural intentions and perceptions-only scores (.31) than with difference scores (.26).

A number of authors have noted that expectations scores are illusory because the most likely response to statements on expectations of service delivery is 'strongly agree' (Carman, 1990; Reynoso and Moores, 1995). In supporting this argument, Babakus and Inhofe (1991) suggest that customers are driven by the "I-have-high-expectations" social norm and this creates a bias towards social desirability. The tendency to rate expectations consistently highly means that perceptions tend to be the dominant contributor in the gap score (Babakus and Boller, 1992) and therefore one can question the usefulness of collecting separate expectations scores (Boulding *et al.* 1993; Brown *et al.* 1993). This concern is borne out by the very high (Majority >6 on a 7-point scale) reported averages for expectations in a number of studies (see Parasuraman *et al.* 1991b; Silvestro, 2005). Van Dyke *et al.* (1997) comment on this problem stating that, because most expectations items are vector attributes (i.e. more is always better), the tendency towards extreme responses increases (i.e. 7 on a 7-point scale) and with that the value of expectations scores.

Further objections to gap-based measurement of service quality centre around the difficulty in defining customers' expectations. In his critique of the SERVQUAL model, Teas (1994) notes that the definition of expectations have been defined in a great many ways: desires, wants, normative expectations, desired service, the level of service a customer hopes to receive, and what a service provider should possess (Parasuraman *et al.* 1985, 1988, 1991b; Zeithaml *et al.* 1993). The variety of definitions for expectations creates a loosely defined conceptualisation of service quality (Cronin and Taylor, 1994). Teas (1993a) argues that a considerable amount of measurement error is caused by the different ways which customers define expectations. Parasuraman *et al.* (1991b, 1994) have responded to the concern regarding the definition of expectations as 'normative expectations of service providers', by redefining expectations as 'the service a consumer would expect from 'excellent service organisations'.

There are also practical concerns regarding the validity of expectations scores that are collected at the same time as perceptions scores (i.e. contemporaneously / *ex post*) or measured when they are not clearly formed (Carman, 1990; Iacobucci, 1994). Kahneman and Miller (1986) suggest that consumers may form their expectations as a result of the service provided rather than prior to it – they call these 'experience-based norms.' Clow and Vorhies (1993) argue that unless expectations data is collected before the service delivery it is of questionable use. They 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 over-inflate expectations. However, Carman (1990) notes the impracticality of collecting expectations and perceptions data at different times. A number of authors have noted the boredom factor of two administrations, one for expectations and the other for perceptions (Bourman and van der Wiele, 1992), and the danger that will compromise data reliability. Finally, Carman (1990) suggests that customer expectations are of less value to both academics and practitioners than the relative importance of the various service quality dimensions. He states that service quality can be defined as follows:

$$Q = \sum I_i (P_i - E_i)$$

where  $I$  is importance,  $P$  is perception and  $E$  is expectation of service attribute  $i$ .

#### ***Alternatives to the Gap-based Measures of Service Quality***

As a result of concerns regarding both the theoretical and practical value of gap scores, Cronin & Taylor (1992, 1994) propose a performance-only approach to service quality measurement. SERVPERF uses the same 22 perception items as SERVQUAL, but does not include the set of expectations statements. Their findings and subsequent research shows that a perceptions-only measure of service quality tends to have better predictive validity than a gap approach.

Parasuraman *et al.* (1994a) accept that performance-only measures of service quality do appear to more accurately predict overall quality. However, they argue that the improvement in predictive accuracy comes at the cost of diagnostic value: ‘SERVQUAL could be superior in terms of pinpointing areas of deficiency within a company’. Dean (1999) concurs with this view and supports the use of gap scores in measuring service quality because of its diagnostic value. This highlights the importance of considering the practical value of a scale when making assessments of scales (Parasuraman *et al.* 1994b). Furthermore, Parasuraman *et al.* (1994b) suggest that direct measures of service quality may suffer from over-inflation of customer service ratings. They argue that evidence from a number of authors shows an upward bias in direct measures of service quality compared with difference-score measures (Peterson and Wilson, 1992; Liljander and Strandvik, 1992; Brown *et al.* 1993).

Carman’s (1990) solution to the problems with the expectations element of SERVQUAL is to suggest that data regarding the perceptions-expectation gap should be

collected in single statements – e.g. ‘The visual appeal of XYZ’s physical facilities is [much better; better; about the same; worse; much worse] than I expected’. He argues that this approach is particularly appropriate when expectations are well established. Babakus and Boller (1992) and Babakus *et al.* 1993) support this notion and argue that, not only would this make the questionnaire easier to complete, it would also reduce the confusion caused by referring to a whole industry in the expectations section and to a specific company in the perceptions section. Finally, Carman (1990) suggests that, ‘in the absence of major change in the service delivery, expectations information might be collected as infrequently as once every third year’ (p49).

### ***Measuring Internal Service***

Internal service focuses on the quality of service ‘provided by distinctive organisational units or the people working in these departments to other units or employees within the organisation’ (Stauss, 1995). The units or employees, termed ‘internal customers’, are similar to external customers in that they have expectations of the service they should receive and make judgements based on these expectations (Bruhn, 2003).

Compared with external service research, there are relatively few academics focused on internal service encounters (Stanley and Wisner, 2001). This is partly a consequence of the marketing background of many service quality academics (Reynoso and Moores, 1995) and the multi-disciplinary nature of internal service (Hallowell *et al.* 1996; Farner *et al.* 2001). Of the work that does exist, the majority is descriptive and does not seek to empirically measure perceptions of service delivery (Frost and Kumar, 2000; Stanley and Wisner, 2001).

Where studies do attempt to measure internal service, the most common approach has been to apply disconfirmation theory through the SERVQUAL tool. Kang *et al.* (2002) argue that, if disconfirmation theory is applicable in both external and internal settings, it is reasonable to assume that SERVQUAL is too. Another justification for applying externally-developed service quality measures to internal settings is that the relationship between external customer and an organisation is just one link in network of interactions, many of which occur within an organisation (Heskett *et al.* 1997; Auty and Long, 1999).

However, there is no consensus on the applicability of SERVQUAL for internal service measurement. Internal and external customers differ in terms of what they consume, the level of choice, and the level of expertise in the goods or services they are purchasing (Finn *et al.* 1996). Whilst similarities exist between the two groups, it should not simply be assumed that the dimensions by which they assess service quality are exactly

the same (Brooks *et al.* 1999). The content validity of SERVQUAL is questioned by the need for additional items and dimensions not incorporated in the original scale, whilst construct validity concerns largely relate to the rejection of the tangibles dimension for internal service (Young and Varble, 1997; Brooks *et al.*, 1999; Auty and Long, 1999; Kuei, 1999). Whilst minor changes are necessary for each research setting (Zeithaml *et al.*, 1990), major alterations bring into question the reliability and validity of what remains. More broadly, there remains no assessment as to the relative merits of measuring internal service based on disconfirmation theory (the gap-approach) as opposed to attitudinal theory (the perceptions-only approach). **Table 1** summarises the SERVQUAL applications in an internal service setting.

#### Take in table 1

The result of such concerns over the theoretical and operational efficacy of SERVQUAL is for some academics to develop specific measures for internal service (Cavinato, 1987; Hendrick & Ruch, 1988; Lewis and Gabrielson, 1995; Lewisohn and Reynoso, 1995; Hallowell *et al.* 1996; Rossler & Hirz, 1996; Marshall *et al.* 1998; Gilbert, 2000; and Bruhn, 2003). The majority of these scales explicitly or implicitly employ attitudinal theory when developing their scales. However, none have empirically assessed how appropriate this approach is.

### 3. Research Objectives

Within the internal service literature, the need to assess the psychometric and practical value of scales based on the disconfirmation model (the gap approach) as opposed to the attitudinal model (the perceptions approach) remains an unresolved challenge. Therefore, the main research objective was to compare two internal service scales – one based on gap scores, the other using perceptions-only scores. Our research focused on the following questions:

- *Is disconfirmation theory transferable to internal service quality?*
- *How reliable and valid is the gap-based scale?*
- *How reliable and valid is the perception-only scale?*
- *What are the benefits and limitations of each in an internal service context?*

In testing these propositions, a service quality measurement instrument was specifically designed to fit the chosen operational context. The internal service involved the provision of e-procurement software and a range of supporting services including training and user support within four UK organisations. Full details of to the methodology used for the development of the measurement instrument are not provided here, as the focus of this paper is not the design of the instrument, but rather analysis and evaluation of the gap versus perception-only approach to measuring internal customer perceived quality.

#### **4. Methodology**

The study used a theoretical sample of organisations ranging in size, budget, implementation strategy, and procurement activity (**Table 2**). The focus was on internal customers (users) of an e-procurement system across the four cases. The research employed both qualitative and quantitative methods to develop and validate a measure of internal service – the *EPQ Scale*.

##### **Take in table 2**

A questionnaire based on items generated in semi-structured interviews was developed consisting of 33 paired-statements relating to each item in the proposed *EPQ Scale*. The first set related to expectations, the second to perceptions. These were anchored 1-7 Likert scales from ‘Strongly Disagree’ to ‘Strongly Agree’. Additionally, a single ‘overall e-procurement quality rating’ was included to assess the predictive power of the proposed scale – a common approach in scale validation (cf. Parasuraman *et al.* 1988; Pitt *et al.* 1995). After piloting, the questionnaire was distributed to all 295 e-procurement users across 4 organisations. The sample-to-variable ratio of 8.3-to-1 exceeds nearly all recommendations (cf. Gorsuch, 1983; Hatcher, 1994; Bryant and Yarnold, 1995; Velicer and Fava, 1998; Hutcheson and Sofroniou, 1999). To ensure a high response rate, contact was made with potential respondents prior to sending our questionnaires. 274 returned usable questionnaires, representing an overall response rate of 92.9%. All data was entered in *SPSS* version 13.01 for statistical analysis. Exploratory factor analysis was employed, with 74.8% of total variance and 68.8% common variance extracted – well above the 60% minimum suggested by Hair *et al.* (1998). Cronbach Alpha, item-to-total correlations,

regression, and analysis of variance were all employed to assess the reliabilities and validities of two scales – one based on gap scores, the other on perceptions-only scores.

## **5. Data Analysis – Gap Approach**

**Table 3** shows the final factor solution for the internal service scale based on gap scores, with details of factor loadings, item-to-total correlations, and alpha coefficients.

**Take in table 3**

### **5.1 Scale Reliability (Gap Approach)**

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 .751 to .954, easily exceeding the recommended cut-off points of .70 and .60 (Nunnally, 1978). These results combined with item-to-total scores (.539 to .903, average .716) indicate a high level of internal consistency between items making up each factor. The overall alpha for the scale is .949.

### **5.2 Scale Validity (Gap Approach)**

The high reliabilities and clear factor structure provide support for trait validity of the *EPQ Scale*. 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). In explicating the user-perceived e-procurement quality construct, a wide range of literature was drawn on and combined with both qualitative and quantitative data. 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 for one another (Carman, 1990). The rules of variable convergence and discrimination hold good for this data. 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.

Predictive validity is derived by examining the predictive power of scale scores on a separate criterion (Flynn *et al.* 1990). **Table 4** illustrates that just under half of variance in the independent measure of internal service – the *OE PQ Rating* – is explained by the average of six factors ( $R^2$  .486). Subsequently, the predictive validity of the factors has been examined using a step-wise multiple regression (**Table 5**). This helps to assess the relative value of each factor to the regression model. The best solution has four factors explaining 54.9% of variance in *OE PQ Ratings*.

**Take in tables 4 and 5**

The statistical power of the regression model is partly determined by the number of independent variables and the significance level chosen (Hair *et al.* 1998). 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. The other factor influencing regression model significance is sample size. Very small samples (<20) only allow the use of simple regression with one independent variable and only very strong relationships detected with certainty. Very large samples (>1000) create problems of over-sensitivity in statistical tests. Hair *et al.* (1998) suggest that, assuming a representative sample, the ratio of observations to independent variables should always be greater than 5-to-1, ideally 20-to-1, and if stepwise regression is applied, closer to 50-to-1. In this research, the ratio of observations to independent variables is 45.7-to-1.

**Table 6** summarizes the results of the analysis of variance (*ANOVA*). The output for the regression displays information about variation accounted for by the four-factor regression solution. Residual output shows information regarding unaccounted variation. A comparison of the regression sum of squares and the residual sum of squares indicates that the model accounts for significant variation in the dependent variable. The F-statistic is the regression mean square (*MSR*) divided by the residual mean square error (*MSE*). The small significance value of this statistic indicates that the *EPQ* factors based on the gap approach do a good job explaining variation in the *OE PQ Rating*.

**Take in table 6**

In examining standardized residuals, 271 of the 274 cases were less than 3 standard deviations away from the mean. This indicates that, in all but three cases, the regression model accurately predicts *OEPQ Ratings*. The normal distribution of residuals (The difference between the observed and predicted value) is further evidence of the scale's predictive validity (**Fig 1**).

**Take in figure 1**

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.

## **6. Data Analysis – Perceptions Approach**

In examining the data from this research one can see the potential extent of the problem of inflated expectations scores (**Table 7**). Of the 30 retained items, none have an expectations mean below 5.5, with an average of 6.43 on a 1-7 Likert scale. The limited variation in expectation scores means perceptions scores are the dominant component for the data set.

**Take in table 7**

Factor analysis of perceptions only scores is undertaken in order to assess the relative merits of the two construct operationalisations. To ensure a fair comparison between the two options, identical choices have been made regarding method selection, factor design, assumption testing, retention of factors, extraction, rotation, interpretation, scale purification, creation of summated scales, and validation. **Table 8** shows the final factor solution for the internal service scale based on perceptions-only scores, with details of factor loadings, item-to-total correlations, and alpha coefficients. The procedure results in an identical six-factor solution. Moreover, all the perceptions variables load on the same factors as their corresponding gap variables. With the exception of one additional item – *order accuracy* – all variables loaded on a single factor. The scale explains 74.76% of *total variance* and 68.75% of *shared variance*, an improvement of 3.85% and 4.75% respectively over the gap approach.

**Take in table 8**

### **6.1 Scale Reliability (Perceptions Approach)**

Internal reliability is indicated by Alpha coefficients which range from .780 to .966 for the six factors and .958 for the entire scale. These coefficients represent a marginal improvement on the gap-approach. In addition, the scale based solely on perceptions scores has slightly higher item-to-total scores than the gap-approach solution.

### **2.2 Scale Validity (Perceptions Approach)**

Content validity can be seen as the same for the two approaches as they produce almost identical scales. In terms of construct validity, the rules of variable convergence and discrimination (Bagozzi, 1981) also hold good for the perceptions-only data. Of the original 33 variables, 29 have high loadings on a single factor. In addition, the scale exhibits high Alphas and high item-to-total scores. Multiple regression analysis was applied to assess the predictive validity of the perceptions-only scale. **Table 9** illustrates that 59% of variance in the *OEPQ Rating* is explained by the average of six factors. This compares with 48.6% of variance explained by the gap-approach. Step-wise multiple regressions suggest that the best solution has five factors explaining 66.5% of variance in *OEPQ Ratings* (**Table 10**). This compares with the four-factor solution explaining 54.9% of variance in the gap-approach.

**Take in tables 9 and 10**

*ANOVA* results for the five-factor regression solution are shown in **table 11**. The residual output for the perceptions-only solution is also lower than for the gap approach – 140.337 compared to 189.593, indicating a lower level of unexplained variance in the data.

**Take in table 11**

In examining standardized residuals, 270 of the 274 cases were less than 3 standard deviations away from the mean. This indicates that, in all but four cases, the regression model based on perception-only data accurately predicts *OEPQ Ratings*. The normal distribution of residuals is further evidence of the scale's predictive validity (**Fig 2**).

**Take in figure 2**

### ***Summary of Analysis***

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 12**). However, it is also clear that in many areas the perceptions-only scale does outperform the gap scale, although the difference in performance is marginal.

Take in table 12

## **7. Discussion**

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

- ***Is disconfirmation theory transferable to internal service quality?***

This research suggests that disconfirmation theory is transferable to internal service quality but practical merit of the gap measure is likely to vary by operational context.

- ***How reliable and valid is the gap-based scale?***

This study confirmed the validity and reliability of the gap-based scale. Gap-based measurement facilitates both an understanding of changes in expectations and perceptions over time. This is particularly important in industries where expectations are poorly understood: this can apply particularly to internal services where there has traditionally been a dearth of research into internal customer perceptions. It is also particularly relevant 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. 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. It is also argued that gap-based measures have diagnostic value, and that difference scores can better pinpoint areas of deficiency within an organisation (Pitt *et al.* 1995, 1997; Dean, 1999).

- ***How reliable and valid is the perception-only scale?***

Whilst the perceptions-only scale only shows marginal improvements in reliability and validity, one must also consider the practical implications of the substantially increased questionnaire length of the gap approach when assessing the two options. There is a clear trade-off here between the data richness and diagnostic value of the paired-statement gap approach, compared with the marginally higher reliability, validity and collection efficiencies gained from single-statement data. In internal service contexts where expectations are perceived as relatively stable, it may be appropriate, as Carman, 1990 advised, to measure expectations separately as infrequently as once every three years. 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). Finally, when the focus of study is on prediction of related constructs, the perceptions-only approach may be most appropriate.

- ***What are the benefits and limitations of each in an internal service context?***

The benefits and limitations of the two approaches are summarized in **table 13**.

**Take in table 13**

### ***An alternative approach***

It has been argued that the choice between adopting a gap-based versus a perceptions-only measure of internal service quality is largely a practical managerial decision which must be based on the operational contingencies regarding the nature of internal customer relationships and the degree of organisational turbulence experienced in a particular context. If the decision is taken not to measure internal expectations separately, there is an alternative approach to the perceptions-only measure which could be considered. It may be better to re-write single statements which capture the perceptions-expectations gap rather than simply using the perceptions half of the original paired statements. Carman (1990), Babakus & Boller (1992), and Reynoso & Moores (1995), advocate this single statement gap-based approach because it embraces the theoretical underpinnings of internal service quality. **Table 14** illustrates how paired-statements could be re-written as single statements.

**Take in table 14**

## 8. Conclusion

The managerial decision as to which option might be most appropriate is contingent on a number of variables which we have identified from the literature, combined with the empirical findings reported in this study. The following summary of the contingencies yields a framework which managers can use in order to base their methodological design decisions. Perceived internal service quality measurement may be operationalised in three ways:

### **Option 1: Paired-statement – gap between expectations and perceptions**

This option may be most appropriate when:

- Internal customer expectations are not known
- Internal customer expectations change significantly over time
- The organisation is undergoing rapid and turbulent change
- Some expectations items may not be vector attributes
- Focus is on the gap between expectations and perceptions
- Questionnaire length is less critical

### **Option 2: Single-statement – perceptions only**

This option may be most appropriate when:

- Expectations are clearly established and stable
- The organisational environment is stable
- All expectations items appear to be vector attributes
- Prediction of related constructs is critical
- Minimal questionnaire length is important
- Potential upward bias of ratings is not critical

### **Option 3: Single-statement – gap between expectations and perceptions**

This option may be most appropriate when:

- Theoretical underpinning of internal service is considered important
- Focus is on the gap between expectations and perceptions
- Prediction of related constructs is important
- Minimal questionnaire length is important
- Potential over-inflation of service ratings is considered critical

Rather than stating that one method of construct measurement is superior, the contingency approach helps researchers and practitioners to operationalise the measurement of internal service quality based on their own objectives and based on the organisational and environmental conditions in which the internal service is operating. There is a clear need to further test this single-statement gap measurement approach in other internal service contexts.

Indeed further empirical testing of all three measurement approaches should facilitate further development of our understanding of the contingencies of internal service quality measurement implementation.

To conclude, in this paper a comparative evaluation has been made of the gap-based and perception-only approach towards the measurement of internal service quality. Both approaches could be justified theoretically, and through empirical testing it was established that both approaches can be operationalised in ways which are reliable and valid. A third approach, based on single-statement gap measures was also identified as a viable alternative approach which might be applied in some internal service contexts. The empirical study, combined with the contributions from the literature, has generated some understanding of the specific organisational and managerial conditions in which the three approaches might be appropriate. The debate as to which approach is better is therefore superceded by a perhaps more productive approach: one which aims to develop a better understanding of the contingencies which determine appropriate implementation.

## 9. References

- Auty, S and Long, G. (1999). 'Tribal Warfare' and gaps affecting internal service quality. *Internal Journal of Service Industry* Vol. 10, 7-18.
- Babakus, E. Boller G. W. (1992). An Empirical Assessment of the SERVQUAL Scale. *Journal of Business Research* 24, 253-268.
- Babakus, E. Inhofe M. (1991). The role of expectations and attribute importance in the measurement of service quality. in Gilly, M.C. et al. (Eds) *Proceedings of the Summer Educators' Conference, American Marketing Association, Chicago, IL.* 142-144.
- Babakus, E. Mangold. W. F. (1992). Adapting the SERVQUAL scale to hospital services: an empirical investigation. *Health Services Research* 26, 767-786.
- Bagozzi, R. O. (1981). Attitudes, intentions, and behaviour: a test of some key hypothesis. *Journal of Personality and Social Psychology* 41, 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* 29, 23-42.
- Boulding, W. Kalra A. Staelin R. Zeithaml V. A. (1993). A dynamic process model of service quality: from expectations to behavioural intentions. *Journal of Marketing Research* 7-27.
- Bouman, M. van der Wiele T. (1992). Measuring Service Quality in the Car Service Industry: Building and Testing an Instrument. *International Journal of Service Industry Management* 3, 4-16.
- Brady, M. K. Cronin J. J. Brand R. R. (2002). Performance-only measurement of service quality: A replication and extension. *Journal of Business Research* 55, 17.

- Brooks, RF, Lings, I, and Botschen, M (1999). Internal Marketing and Customer Driven Wavefronts. *The Service Industries Journal* October, 49-67.
- Brown, T. J. Churchill G. A. J. Peter J. P. (1993). Research Note: Improving the Measurement of Service Quality. *Journal of Retailing* 69, 127-139.
- Bruhn, M (2003). Internal Service Barometers. *European Journal of Marketing* 37, 1189-1204.
- Bryant., Yarnold. (1995). Principal components analysis and exploratory and confirmatory factor analysis. In *Grimm & Yarnold, Reading and understanding multivariate analysis. American Psychological Association Books.*
- Buttle, F (1996). SERVQUAL: Review, critique, research agenda. *European Journal of Marketing* 30, 8-33.
- Cadotte, E, Woodruff, R, and Jenkins, R. (1987). Expectations and Norms in Models of Consumer Satisfaction. *Journal of Marketing Research* 24, 305-314.
- Carman, J. M. (1990). Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions. *Journal of Retailing* 66, 33-55.
- Caruana, A and Pitt, L (1997). INTQUAL - an internal measure of service quality and the link between service quality and business. *European Journal of Marketing* 31, 604.
- Cavinato, J (1987). Purchasing Performance: What Makes The Magic? *Journal of Purchasing and Materials Management* 23, 10-16.
- Chaston, I. (1994). Internal customer management and service gaps within the UK manufacturing sector. *International Journal of Operations & Production Management* 14, 45-57.
- Chaston, I (1995). A typology for evaluating branch-level perceptions of internal customer management processes within the UK clearing banks. *The Service Industries Journal* 15, 332-350.
- Churchill, G. A. Jr. (1979). A Paradigm for Developing Better Measures of Marketing Constructs. *Journal of Marketing Research* 16, 64-73.
- Clow, K. E. Vorhie D. E. (1993). Building a competitive advantage for service firms . *Journal of Services Marketing* 7, 22-23.
- Cronin, J and Taylor, S. (1994). SERVPERF versus SERVQUAL: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality. *Journal of Marketing* 58, 125-131.
- Cronin, J. J and Taylor, S. (1992). Measuring Service Quality: A Reexamination and Extension. *Journal of Marketing* 56, 55-68.
- Dean, A. (1999). The applicability of SERVQUAL in different health care environments. *Health Marketing Quarterly* 16, 1-21.
- Farner, S, Luthans, F, and Sommer, S. (2001). An empirical assessment of internal customer service. *Managing Service Quality* 11, 350-358.
- Finn, D, Baker, J, Marshall, G, and Anderson, R (1996). Total Quality Management and Internal Customers: Measuring Internal Service Quality.
- 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, 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, 358-377.

- Frost, F and Kumar, M (2001). Service Quality between internal customers and internal suppliers in an international airline. *The International Journal of Quality and Reliability Management* 18, 371-386.
- Ganesh, J, Arnold, M, and Reynolds, K. (2000). Understanding the customer base of service providers: An examination of the difference between switchers and stayers. *Journal of Marketing* 64, 65-87.
- Gilbert, G R. (2000). Measuring internal customer satisfaction. *Managing Service Quality* Vol. 10, 178-185.
- Gorsuch, R. L. (1983). Factor Analysis. *Hillside, NJ: Lawrence Erlbaum*.
- Grönroos, C. (1982). Strategic Management and Marketing in the Service Sector. *Helsingfors: Swedish School of Economics and Business Administration*.
- Hair, J. F. Jr. Anderson R. E. Tatham R. L. Black W. C. (1998). Multivariate Analysis 5th Edition. *Prentice-Hall*.
- Hallowell, R, Schlesinger, L, and Zornitsky, J (1996). Internal Service Quality, Customer and Job Satisfaction: Linkages and Implications for Management. *Human Resource Planning* 19, 20-31.
- Hatcher, L. (1994). A step-by-step approach to using the SAS system for factor analysis and structural equation modeling. *Cary, NC. SAS Institute. Focus on the CALIS procedure*.
- Hendrick, T and Ruch, W (1988). Determining Performance Appraisal Criteria for Buyers. *Journal of Purchasing and Materials Management* 24, 18-26.
- Heskett, J. L. Sasser W. E. 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* 8, 229-242.
- Hutcheson, G. Sofroniou N. (1999). The multivariate social scientist: Introductory statistics using generalized linear models. *Thousand Oaks, CA. Sage Publications*.
- Iacobucci, D. Grayson K. A. 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. 3.
- Kahneman D., Miller K. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychol. Reviews* 93, 136-153.
- Kang, Gi-Du, James, J, and Alexandris, K. (2002). Measurement of internal service quality: application if the SERVQUAL battery to internal service quality. *Managing Service Quality* 12, 278-291.
- Lewis, B. R. Gabrielson G. O. S. (1995). An Intra-organisational Approach towards the Implementation of Service Quality Management. *Manchester School of Management*.
- Lewis, R. C. Booms B. H. (1983). The Marketing Aspects of Service Quality in *Emerging Perspectives on Services Marketing*, Berry, L., Shostack, G., Upah, G. *Chicago: American Marketing*.
- Lewisohn, C and Reynoso, J (1995). Improving health service quality from within: The case of United Leeds Teaching Hospitals NHS Trust. *International Journal of Health Care Quality* 8, 18-21.
- Liljander, V. Strandvik. T. (1992). The Relationship between Service Quality, Satisfaction and Intentions. *Proceedings in the 2nd Workshop on Quality Management in Services*, Jos Lemmink and Paul Kunst (eds.) 77-98.
- Lings, I. (2000). Internal Marketing and supply chain management . *Journal of Services Marketing* Vol. 14, 27-41.

- Lings, I. Brooks R. (1998). Implementing and measuring the effectiveness of internal marketing. *Journal of Marketing Management* 14, 325-351.
- Marshall, G, Baker, J, and Finn, D (1998). Exploring internal customer service quality. *Journal of Business & Industrial Marketing* Vol 13, pp381-392.
- Nunnally, J. C. (1978). Psychometric Theory 2nd Ed. *McGraw-Hill, New York, NY*.
- Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research* 17, 460-469.
- Oliver, R. L. (1999). Whence consumer loyalty. *Journal of Marketing* 33-44.
- Parasuraman, A, Berry, L. L, and Zeithaml, V. (1991a). Perceived service quality as a customer-based performance measure: An empirical examination on organizational barriers using an extended service quality model. *Human Resource Management* 30, 335-364.
- Parasuraman, A, Zeithaml, V, and Berry, L (1985). A conceptual model of service quality and its implications for future research. *Journal of Services Marketing* 49, 41-50.
- 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* 58, 111-124.
- 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* 64, 12-40.
- Parasuraman, A. Berry L. Zeithaml V (1991b). Refinement and Reassessment of the SERVQUAL scale. *Journal of Retailing* 67, 420-450.
- Parasuraman, A. Zeithaml V. Berry L. (1994b). Alternative Scales for measuring service quality: A Comparative Assessment Based on Psychometric and Diagnostic Criteria. *Journal of Retailing* 70, 201-230.
- Parasuraman, A. Zeithaml V. Berry L. (1993). Research note: more on improving service quality measurement. *Journal of Retailing* 69, 140-147.
- Peter, J. P. Churchill G. A. Jr. Brown T. J. (1993). Caution in the Use of Difference Scores in Consumer Research. *Journal of Consumer Research* 19, 655-662.
- Peterson, R. A. Wilson W. R. (1992). Measuring Customer Satisfaction: Fact and Artifact. *The Journal of the Academy of Marketing Science* 20, 61-71.
- Pitt, L. F, Kavan, R. T, and Bruce, C. (1995). Service quality: A measure of information systems effectiveness. *MIS Quarterly* 19, 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* 21, 209-222.
- 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* 32, 37-43.
- Sekaran, U. (2003). Research Methods for Business: A Skill Building Approach 4th Edition. *Wiley & Sons, Inc. U.S.*
- Silvestro, R. (2005). Applying Gap Analysis in the Health Service to Inform the Service Improvement Agenda. *The International Journal of Quality and Reliability Management* 22, 215-233.
- Smith, A. M. (1995). Measuring service quality: is SERVQUAL now redundant? *Journal of Marketing Management* 11, 257-276.

- Stanley, L and Wisner, J. (1998). Internal Service Quality in Purchasing: An Empirical Study. *International Journal of Purchasing and Materials Management* Summer 34, 50-60.
- Stanley, L. L and Wisner, J. D. (2001). Service Quality along the supply chain: implications for purchasing. *Journal of Operations Management* 287-306.
- Stanley, L. L. Wisner J. D. (2002). The determinants of service quality: issues for purchasing. *European Journal of Purchasing and Supply Management* 8, 97-109.
- Stauss, B (1995). Internal services: Classification and quality management. *International Journal of Service Industry Management* Vol. 6, 62-79.
- Teas, K. R. (1993b). Consumer expectations and the measurement of perceived service quality. *Journal of Professional Services Marketing* 8, 33-53.
- Teas, K. R. (1994). Expectations as a Comparison Standard in Measuring Service Quality: An Assessment of a Reassessment. *Journal of Marketing* 58, 1325-139.
- Teas, K. R. (1993a). Expectations, Performance Evaluation, and Consumers' Perceptions of Quality. *Journal of Marketing* 57, 18-34.
- 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* 21, 195-209.
- Velicer, W. F. & Fava U. (1998). The effects of variable and subject sampling on factor pattern recovery. *Psychological Methods* 3, 231-251.
- 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* 33, 36-41.
- Zeithaml, V. Berry L. Parasuraman A. (1993). The nature and determinants of customer expectations of service. *Journal of the Academy of Marketing Science* 21, 1-12.
- Zeithaml, V. Parasuraman A. Berry L. (1990). Delivering Quality Service: Balancing Customer Perceptions and Expectations. *Free Press, New York, NY*.

## 10. Tables & Figures

**Table 1. *SERVQUAL* Applications in *ISQ* Literature**

Study	Instrument	Analysis	Factor Structure	Assessment of Gap- and Single-Statement?
Chaston '94, '95	36 items based on <i>SERVQUAL</i>	Descriptive analysis of survey data	5 factors, plus <i>Proactive Decision-Making</i>	No
Boshoff & Mels '95	9 items from <i>SERVQUAL</i>	Path Analysis	1 factor: ISQ is uni-dimensional	No
Reynoso & Moores '95	45 items from <i>SERVQUAL</i> + qualitative work	<i>PCA</i> by oblique rotation	10 factors, similar to PZB '85	No
Caruana & Pitt '97	34 items from <i>SERVQUAL</i> + focus groups	<i>PCA</i> by varimax rotation	2 factors comprised of 17 items	No
Hill & McCrory '97	24 items 'guided' by <i>SERVQUAL</i>	Descriptive analysis of survey data	5 factors, new names	No
Young & Varble '97	22 original items	Descriptive analysis of survey data	4 factors: <i>Tangibles</i> unimportant	No
Lings, Brooks, Botschen '98, '99, '00	10 PZB ('85) factors	Descriptive analysis of case data	8 of 10 PZB factors, plus 3 new factors	No
Stanley & Wisner '98, '01, '02	12 items based on <i>SERVQUAL</i>	Descriptive analysis of survey data	9 factors in '02 study	No
Auty & Long '99	22 original items	Mix in range of MBA studies	10 PZB ('85) factors more applicable to ISQ	No
Kuei '99	18 items (no <i>Tangibles</i> )	Cluster analysis		4 factors: <i>Tangibles</i> unimportant
White & Ruddall '99	21 items + 6 ISQ-specific	<i>PCA</i>		5 factors, but items do not load as hypothesised
Frost & Kumar '00, '01	22 items + 2	<i>PAF</i> by oblique rotation		5 factors
Kang <i>et al.</i> '02	22 original items	<i>CFA</i>		5-factor model

*PAF* – Principal Axis Factor Analysis

*PCA* – Principal Components Factor Analysis

*CFA* – Confirmatory Factor Analysis

**Table 2. General Characteristics of Cases**

	Case 1	Case 2	Case 3	Case 4
<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 (G&amp;S)</b>	£600 million	£16 million	£6 million	£15 million
<b>Requisitions P/A</b>	150,000	4000	2000	2900
<b>Active Suppliers</b>	13,000	2500	800	2300
<b>Previous Procurement System</b>	Mix	Mix	Paper	Paper
<b>E-Procurement Start Data</b>	Jan 03	Dec 03	Aug 03	Oct 03
<b>Project Team</b>	Procurement	Procurement	Finance	Procurement
<b>Roll-out Strategy</b>	Commodity	Department	Department	Commodity
<b>System Users</b>	156	44	41	54
<b>Departments using E-Procurement</b>	13 of 15	8 of 9	4 of 4	11 of 11
<b>Level of <i>FMS</i> Integration</b>	Extensive	Limited	None	Limited
<b>Use of Reporting Functionality</b>	High	High	Low	Medium

**Table 3 Perceived EPQ Factor Solution (Gap Approach)**

<b>Variable</b>	<b>Item-to-total</b>	<b>Professionalism Alpha .954</b>	<b>Processing Alpha .897</b>	<b>Training Alpha .919</b>	<b>Specification Alpha .818</b>	<b>Content Alpha .796</b>	<b>Usability Alpha .751</b>
support availability	.807	.830					
support reliability	.818	.784					
support responsiveness	.869	.899					
support knowledge	.840	.822					
support flexibility	.791	.710					
problem resolution	.824	.757					
confidentiality	.817	.829					
friendliness	.763	.867					
concern shown	.793	.919					
order processing speed	.721		.664				
ease of authorisation	.644		.547				
orders to supplier speed	.744		.901				
order lead-time	.756		.807				
processing complex orders	.608		.490				
on-time delivery	.724		.805				
order accuracy	.636		.693				
system security	.574		.567				
timely training	.859			.888			
appropriate training	.903			.982			
information provision	.755			.654			
FMS integration	.599				.666		
invoice reconciliation	.692				.644		
system configurability	.592				.486		
reporting capability	.674				.719		
loaded suppliers	.666					.738	
loaded catalogues	.689					.870	
ease of search	.571					.473	
system availability	.539						.409
screen loading speed	.639						.734
ease of navigation	.565						.625

**Table 4. Linear Regression: EPQ Score to OEPQ Rating (Gap Approach)**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1(a)	.698(a)	.487	.486	.897

a Predictors: (Constant), EPQ

b Dependent Variable: OEPQ Rating

**Table 5. Stepwise Regression: EPQ Factors to OEPQ Rating (Gap Approach)**

Model (e)	R	R Square	Adjusted R Square	Std. Error of the Estimate
1(a)	.669(a)	.448	.446	.931
2 (b)	.723(b)	.523	.520	.866
3 (c)	.736(c)	.542	.537	.851
4 (d)	.745(d)	.556	.549	.840

d Predictors: (Constant), Professionalism, Processing, Training, Specification

e Dependent Variable: OEPQ Rating

**Table 6. EPQ Regressions ANOVA (Gap Approach)**

	Sum of Squares	df	Mean Square	F	Sig.
Regression (a)(b)	237.155	4	59.289	84.121	.000(d)
Residual	189.593	269	.705		
Total	426.748	273			

a Predictors: (Constant), Professionalism, Processing, Training, Specification

b Dependent Variable: OEPQ Rating

**Table 7. EPQ Expectations**

Item	Mean	Item	Mean
(E)* system security	6.73	(E) timely training	6.53
(E) orders to suppliers	6.72	(E) on-time delivery	6.46
(E) system navigation	6.69	(E) order lead-time	6.42
(E) ease of authorisation	6.69	(E) friendliness	6.38
(E) order processing	6.69	(E) confidentiality	6.35
(E) appropriate training	6.65	(E) support availability	6.33
(E) support reliability	6.62	(E) support flexibility	6.30
(E) invoice reconciliation	6.61	(E) FMS integration	6.29
(E) system availability	6.59	(E) processing service orders	6.29
(E) ease of search	6.56	(E) information provision	6.27
(E) problem resolution	6.55	(E) loaded suppliers	6.15
(E) order accuracy	6.54	(E) concern shown	6.13
(E) screen loading	6.54	(E) reporting capability	6.09
(E) knowledge	6.54	(E) system configurability	5.88
(E) support responsiveness	6.53	(E) loaded catalogues	5.69

**Table 8. Perceived EPQ Factor Solution (Perceptions Approach)**

Variable	Item-to-total	Professionalism Alpha .966	Processing Alpha .910	Training Alpha .930	Specification Alpha .846	Content Alpha .862	Usability Alpha .790
support availability	.838	.808					
support reliability	.859	.883					
support responsiveness	.905	.916					
support knowledge	.891	.893					
support flexibility	.809	.749					
problem resolution	.856	.783					
confidentiality	.864	.851					
friendliness	.819	.923					
concern shown	.858	.956					
order processing speed	.745		.738				
ease of authorisation	.669		.565				
orders to supplier speed	.812		.972				
order lead-time	.765		.754				
processing complex orders	.649		.428				
on-time delivery	.766		.689				
order accuracy	.661		<b>.494</b>		<b>.421</b>		
system security	.608		.530				
timely training	.866			.902			
appropriate training	.916			.989			
information provision	.795			.678			
FMS integration	.651				.605		
invoice reconciliation	.735				.680		
system configurability	.657				.619		
reporting capability	.692				.662		
loaded suppliers	.738					.853	
loaded catalogues	.785					.920	
ease of search	.685					.705	
system availability	.605						.547
screen loading speed	.647						.860
ease of navigation	.600						.493

**Table 9. Linear Regression: EPQ Score to OEPQ Rating (Perceptions Approach)**

Model (b)	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.769(a)	.592	.590	.800

a Predictors: (Constant), PEPQ

b Dependent Variable: OEPQ Rating

**Table 10. Stepwise Regression: EPQ Factors to OEPQ Rating (Perceptions Approach)**

Model (f)	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.760(a)	.578	.576	.814
2	.798(b)	.637	.635	.756
3	.810(c)	.657	.653	.737
4	.816(d)	.666	.661	.728
5	.819(e)	.671	.665	.724

e Predictors: (Constant), PProfessionalism, PSpecification, PTraining, PProcessing, PContent

f Dependent Variable: OEPQ Rating

**Table 11. EPQ ANOVA (Perceptions Approach)**

	Sum of Squares	Df	Mean Square	F	Sig.
Regression (a) (b)	286.411	5	57.282	109.391	.000(e)
Residual	140.337	268	.524		
Total	426.748	273			

a Predictors: (Constant), PProfessionalism, PSpecification, PTraining, PProcessing, PContent

b Dependent Variable: OEPQ Rating

**Table 12. EPQ Scale Validation: Gap Approach vs. Perceptions Approach**

Gap Approach Scale		Perceptions Approach Scale
	<b>Reliability</b>	
.751 - .954	Factor Alpha Range	.780 - .966
.949	Scale Alpha	.958
.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> .486	Regression – EPQ Score to OEPQ	Adjusted R <sup>2</sup> .590
Adjusted R <sup>2</sup> .549	Regression – Stepwise Factors to OEPQ	Adjusted R <sup>2</sup> .665
189.59 from 426.75	ANOVA of Residuals (Unaccounted Variation)	140.34 from 426.75

**Table 13: Evaluation of Gap-based versus Perceptions-only methodologies**

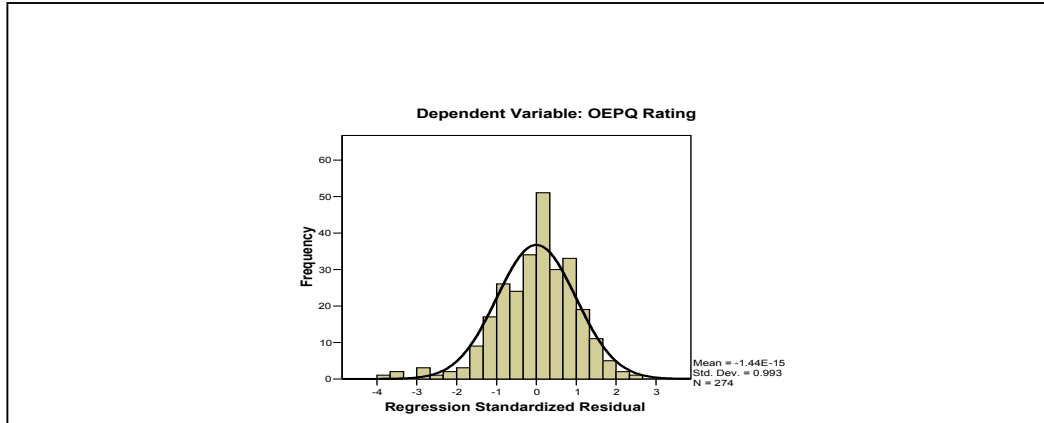
	<b>Gap-based measure</b>	<b>Perceptions-only measure</b>
<b>BENEFITS</b>	<ul style="list-style-type: none"> <li>Valid and reliable</li> <li>Data richness</li> <li>Improved understanding of expectations</li> <li>Increased diagnostic value: effective in identifying improvement priorities</li> </ul>	<ul style="list-style-type: none"> <li>Marginally increased reliability and validity</li> <li>Marginally 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>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> </ul>

**Table 14. EPQ Question Modification**

	<b>Well below my expectations</b>				<b>Well above my expectations</b>		
The availability of the system.	1	2	3	4	5	6	7
The speed of screen loading.	1	2	3	4	5	6	7

**Summary**

**Figure 1. Histogram of *EPQ* Residuals (Gap Approach)**



**Figure 2. Histogram of Single-Statement *EPQ* Residuals (Perceptions Approach)**

