ODLPERF: AN INSTRUMENT FOR MEASURING SERVICE QUALITY IN AN OPEN AND DISTANCE LEARNING (ODL) INSTITUTION

Ramli Bahroom
Corporate Planning Unit
Open University Malaysia
Jalan Tun Ismail, 50480 Kuala Lumpur, Malaysia
Tel: +60327742710
Fax: +60326978816
ramli@oum.edu.my

Latifah Abdol Latif
Centre for Student Management
Open University Malaysia
Jalan Tun Ismail, 50480 Kuala Lumpur, Malaysia
Tel: +60327742112
Fax: +60326978816
latifah@oum.edu.my

Ng Man San
Centre for Student Management
Open University Malaysia
Jalan Tun Ismail, 50480 Kuala Lumpur, Malaysia
Tel: +60327742112
Fax: +60326978816
mslymy@oum.edu.my

Abstract

Service quality is important in higher education institutions, particularly in an open and distance learning (ODL) environment, where learner attrition rates are generally higher than that of their conventional counterparts. This has led to a rigorous quest for an effective service quality measurement instrument to provide the institutions with the appropriate information that would help meet the expectations of the students in the provision of quality study programmes, teaching and learning and other educational services. This paper reports on a study that was carried out at Open University Malaysia (OUM) to develop an instrument referred to as ODLPERF for measuring service quality for ODL institutions. The study employed a 29-item Importance-Performance Survey questionnaire that was administered randomly to 2,491 under-graduate learners of OUM in 2008 of which 894 were finally returned and used in this study. Using the AMOS 16.0 computer package and applying the methods of exploratory and confirmatory factor analyses, the study identified four factors, namely tangibles, reliability, assurance and empathy as the key dimensions of service quality. The four-factor model was analyzed based on the maximization likelihood (ML) estimation method. The recommended goodness-of-fit indices of the model were found to be within tolerable ranges (RMSEA =0.06, GFI=0.94, AGFI=0.92, CFI=0.95), suggesting that the model provides a close fit to the data. The scales were also found to show acceptable internal consistency reliability with both

Cronbach's Alpha and construct reliability ranging from 0.8 to 0.9, above the recommended threshold of 0.7. To provide evidence of high convergent validity, the standardized loading estimates for each of the four dimensions were found to be above the recommended threshold of 0.5. The average variance extracted (AVEs) were also found to be within the acceptable range of 0.5-0.6. The discriminant validity for the four dimensions was evident as the AVEs were found to be equal to or greater than the squared correlations between any two dimensions. Collectively, the above tests provide reasonable support for reliability and validity of the service quality construct. This suggests that the proposed ODLPERF is a valid and reliable instrument to measure service quality in an ODL institution.

Introduction

The world of higher education is undergoing a rapid change in the last couple of decades. With the advent of the Internet, learners are becoming increasingly complex and their demands of the education services from higher education institutions (HEIs) are more diverse. To add to that, today's globalised world has led to immense and intense competition among HEIs not seen before. However, on a more positive note, the Internet and globalization have opened a lot of opportunities for those institutions that are willing to provide the services required by this new breed of learners. Of course, in this regard, the quality of these services has to commensurate with their provision. Thus, service quality is important in HEIs, and this is more so in an open and distance learning (ODL) environment, where learner attrition rates are generally higher than that of their conventional counterparts.

This has led to a rigorous quest for an effective service quality measurement instrument to provide the institutions with the appropriate information that would help meet the expectations of the students in the provision of quality study programmes, teaching and learning and other educational services. This paper reports on a study that was carried out at Open University Malaysia (OUM) to develop an instrument referred to as ODLPERF (ODL PERFormance) for measuring service quality for ODL institutions.

It has been well accepted that service quality is a fundamental and essential aspect of service provision as well as sustainability of an institution, business of otherwise. The issue is how do we measure service quality? There had been many attempts to measure it but unfortunately, the issue has not been adequately resolved. Parasuraman, *et al.*, (1988) presented the first set of instruments to attempt to measure service quality, which came to be widely known as SERVQUAL. While it does present a viable instrument for the industries that it covered, diverse studies using these scales have demonstrated the existence of difficulties resulting from the conceptual or theoretical component as much as from the empirical component. As a result, many researchers agree that performance-based as opposed to perception-based measure explains more of the variance in the overall measure of service quality (Oliver, 1989; Bolton and Drew, 1991a, b; Cronin and Taylor, 1992; Boulding et al., 1993; Quester et al., 1995). Cronin and Taylor (1992) provided an alternative instrument and called it SERVPERF which focus only on the performance component and leave out the perception component. SERVPERF was found to be better than the SERVQUAL scales (Cronin and Taylor, 1992; Parasuraman et al., 1994; Quester et al., 1995; Llusar and Zornoza, 2000).

Subsequently, by conducting a comprehensive comparative analysis of SERVPERF and a newly introduced HEDPERF, Firdaus (2005 and 2006) found that the former is a generic scale and may not be adequate to measure service quality performance of HEIs and the latter is more appropriate, instead. Firdaus argued that with increasing competition, HEIs need to be learner-centred and the measurement of its service quality has to reflect this.

Firdaus had tested his model based on *unidimensionality, reliability, validity, explained variance and biasedness* and found that HEDPERF is superior to SERVPERF. However, HEDPERF itself, while targeted for HEIs, is still a generic instrument, in that it does not differentiate between different types of HEIs that are in existence today. One of these types is the ODL institutions. ODL institutions are unique in itself, particularly in terms of their delivery modes, degree of flexibility and type of learners. A generic instrument for conventional HEIs as proposed by Firdaus may not be suitable and appropriate to measure its service quality. This paper reports on a study that was carried out at Open University Malaysia (OUM) to develop an instrument referred to as ODLPERF for measuring service quality specifically for ODL institutions. At the onset, it must be emphasized that this exercise is exploratory in nature and thus may exclude several rigorous analytical procedures.

Methodology

The proposed instrument to be used as a performance measure for ODL institutions is a measurement scale which was developed based on an extensive literature review related to service quality, service performance and student satisfaction based on Importance-Satisfaction surveys conducted by OUM since 2004. A total of 70 service-items were initially identified from the literature review and out of that, 29 items were finally selected for use together with questions relating to socio-economic variables to prepare the draft questionnaire. The content of the draft questionnaire was then reviewed and validated by two experts in the area of learner management and strategic planning for ODL institutions before it is finalized and administered on a full-scale basis to collect the data in the survey.

The final questionnaire consists of three sections, namely, A, B and C. Section A contained 18 questions pertaining to respondents' profile. Section B contained 29 items related to different aspects of ODL institution's service offerings. The performance-based items in Section B were presented as statements in the questionnaire, with the same rating scale used throughout, that is, a 7-point, Likert-type scale that ranging from 1 = "strongly disagree" to 7 = "strongly agree". Finally, in Section C, respondents were also asked to provide an overall evaluation of the *quality of service, satisfaction level* and *intention to complete study programme* also using a similar 7-point Likert-type scale.

The sample size used in this survey was 2,491 undergraduate learners distributed over 53 learning centres and spread over 32 study programmes of OUM. The survey was conducted in June 2008. A total of 894 respondents distributed across all the above-mentioned learning centres and study programmes questionnaires returned the questionnaires.

Results

Factor Analysis

An exploratory factor analysis (EFA) was undertaken to identify the dimensional structure of service quality within OUM, an ODL institution. One critical assumption underlying the appropriateness of EFA is to ensure that the data matrix has sufficient correlations to justify its application (Hair *et al.*, 1995). A first step is visual examination of the correlations, identifying those that are statistically significant. Inspection of the correlation matrix reveals that practically all correlations are significant at $p = 0.01$, and this certainly provides an excellent basis for factor analysis.

The next step involves assessing the overall significance of the correlation matrix with Bartlett test of sphericity, which provides the statistical probability that the correlation matrix has significant correlations among at least some of the variables. The results were significant with a Chi-Square (29, $n = 894$) $= 6.616$ ($p = 0.00$), a clear indication of suitability for factor analysis. Finally, Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was computed to quantify the degree of inter-correlations among the variables and the results indicate an index of 0.92, a sign of adequacy for factor analysis (Kaiser, 1970). As for the adequacy of the sample size, there is a 31-to-1 ratio of observations to variables in this study, which falls within acceptable limits.

In order to gain a better understanding of the factor structure, all the 29 items in Section B of the questionnaire were subjected to a factor analysis utilizing the principal components procedure, which was followed by a varimax rotation. The decision to include a variable in a factor was based on factor loadings greater than ± 0.5 (Comrey & Lee (1992, as cited in Lawrence Meyers et al.,(2006); Hair et al., 1995), and all factors whose eigenvalues were greater than 1.0 were retained in the factor solution (Tabachnick and Fidell, 1989). Moreover, the choice regarding factor loadings greater than ± 0.5 was not based on any mathematical proposition but relates more to practical significance (Firdaus, 2005).

The variable's communality, which represents the amount of variance accounted for by the factor solution for each variable was also assessed to ensure acceptable levels of explanation. The results show that communalities in 14 service-items are above 0.50, which are '. . . adequate for having sufficient explanation and are thus retained (Hair *et al.*, 1995, p. 387).

Table 1 shows the results of the factor analysis in terms of factor name, the variables loading on each factor and the variance explained by each factor.

The four factors identified in Table 1 are described as follows:

Factor 1: *Tangibles*. This factor consists of items that are essential to enable students fulfil their study obligations, and it relates to facilities and buildings of the ODL institution.

Factor 2: *Reliability*. The factor comprises of items that describe the degree of dependability that learners can expect out of the academic staff and services.

Factor 3: *Assurance*. This factor is loaded with items that suggest the importance of ODL institutions in meeting learners' expectations in terms of its academic outcomes.

Factor 4: *Empathy*. This factor consists of items that relate to such issues as approachability, ease of contact, availability, caring and convenience.

Table 1: Results of Factor Analysis (Factor Loadings)

| Variables	Factors			
	Tangibles (1)	Reliability (2)	Assurance (3)	Empathy (4)
c25: Quality exam & assignment questions			0.727	
c24: Strict exam invigilation procedures			0.690	
c23: Quality programmes			0.666	
c21: Knowledgeable and competent facilitators/tutors		0.737		
c14: Providing prompt feedback on assignments		0.657		
c15: Providing prompt feedback on online forum discussions		0.633		
c9: Tutorials conducted according to time-table		0.610		
c8: Staff being sympathetic and reassuring				0.792
c13: Attending to enquiries				0.785
c19: Feeling confident with staff				0.781
c17: Staff always willing to help				0.781
c1: Up-to-date T&L facilities	0.790			
c2: Online learning platform	0.744			
c3: Appealing appearance of learning Centre	0.740			
Cronbach's alpha	*0.78*	*0.81*	*0.82*	*0.88*
Cumulative % of variance				*62.9*

Normality test

To test for normality, the normal Q-Q plots of all the four dimensions were obtained as shown in Figure 1 below. As can be seen from the figure, the plots are almost linear thus implying that the data are multivariate normal.

Confirmatory factor analysis

Table 2 presents the measures of goodness-of-fit of the model for the entire sample. The results indicated an acceptable fit for the four-factor model. The overall fit of the model to the data was evaluated in various ways. Specifically, an exact fit of a model is indicated when the p-value for chi-square is above a certain value (usually set to $P > 0.05$) as well as indicated by

other goodness-of-fit measures. While chi-square is sensitive to sample size and tends to be significant in large samples, a relative likelihood ratio between a chi-square and its degrees of freedom was used. According to Eisen et al. (1999), a relative likelihood ratio of 5.0 or less was considered an acceptable fit. In this study, the relative likelihood ratio was found to be 4.5 (Table 2).

Figure 1: Normal Q-Q Plot by Dimension
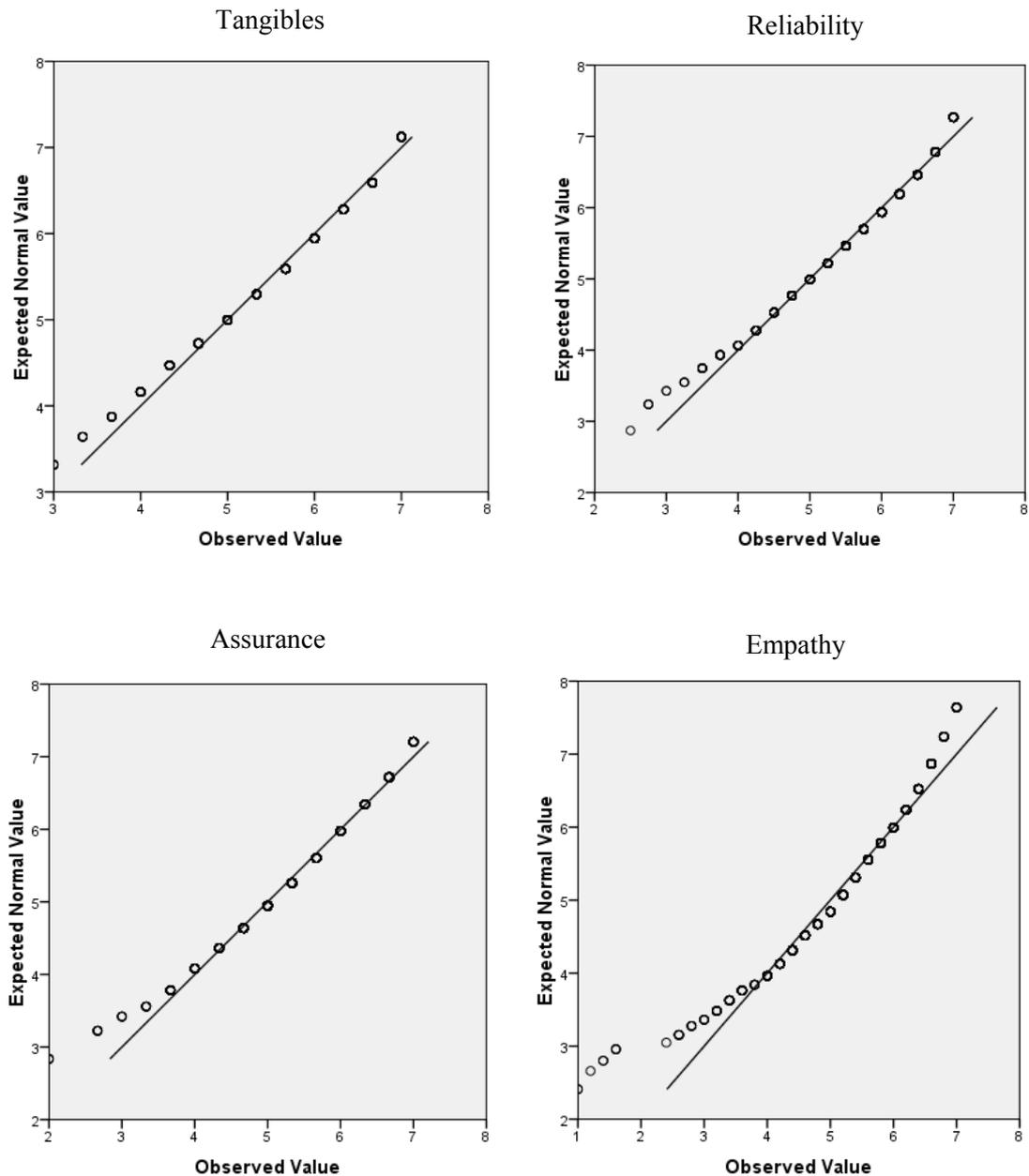
A number of goodness-of-fit measures were used to eliminate or reduce the dependence on sample size for the purpose of confirming whether the four-factor model fits the data. For this purpose, AMOS 16.0 together with the technique of Confirmatory Factor Analysis (CFA) provides many fit indices. These indices include the goodness-of-fit index (GFI), the adjusted

goodness-of-fit index (AGFI), incremental fit indices which include the comparative fit index (CFI), and Normed Fit Index (NFI), and Tucker Lewis Index (TLI). By definition, these indices have values ranging between 0 and 1, with higher values above 0.90 indicating a better fit. In addition, root mean squared error of approximation (RMSEA), the Parsimonious Fit Indices (Parsimonious goodness-of-fit Index (PGFI) and Parsimonious Normed Fit Index (PNFI)) were also computed and evaluated.

Table 2 shows the indices for the four-factor model are within acceptable cut-off values, implying that there is an evidence of unidimensionality for the scales (Bryne, 1994). The root mean squared error of approximation (RMSEA), which is the measure of the discrepancy per degree of freedom, was also reported and its value of 0.06 indicated a fair fit to the data (Kelloway, 1998; Chow *et al.*, 2001). From these results, it appears to indicate that the four-factor model generated in this study fits well and represents a reasonably close approximation to the population.

Table 2: Goodness-of-fit indices for the four-factor model

Statistics	Observed Values	Cut-off values
A. Absolute Fit Indices		
1. Chi-square/degree of freedom	4.50	2 to 5
2. Goodness-of-fit index (GFI)	0.94	> 0.90
3. Adjusted goodness-of-fit index (AGFI)	0.92	> 0.90
4. Root mean square error of approximation (RMSEA)	0.06	< 0.07
B. Incremental Fit Indices		
5. Comparative fit index (CFI)	0.95	> 0.90
6. Tucker Lewis Index (TLI)	0.94	> 0.90
7. Normed Fit Index (NFI)	0.94	> 0.90
C. Parsimononious Fit Indices		
8. Parsimonious Normed Fit index (PNFI)	0.77	> 0.50
9. Parsimonious goodness-of-fit index (PGFI)	0.78	> 0.50

Reliability test

The reliability of the composite score should be assessed after unidimensionality has been acceptably established. In this study, Cronbach's alpha was computed to test for the internal consistency estimates and is shown for each of the dimensions in Table 3. An alpha value of 0.70 and above is considered to be the criteria for demonstrating internal consistency of new scales and established scales respectively (Nunnally, 1988). All the alpha values meet the required prerequisite, thereby demonstrating that all the four dimensions are internally consistent and reliable.

Validity test

Once unidimensionality is established and internal consistency estimates of reliability shows satisfactory values, the next step involves assessing the validity of the dimensions. Validity is the extent to which a measure or set of measures correctly represents the concept and scope of study. For the purpose of validating the four service quality factors dimensions, the following validity tests, namely, criterion validity, discriminant validity and convergent validity were conducted (Table 4). Service quality was significantly correlated with overall quality and overall satisfaction, suggesting criterion validity was satisfied. The average variance extracted (AVE) was found to be in the range of 0.53 to 0.60 which is greater than the square of correlation between any two dimensions, implying there was discriminant validity. Finally, the standardized loadings for each dimension were all above the value of 0.5, indicating that there was convergent validity for each dimension. Convergent validity was also tested by evaluating the AVEs which are above 0.5, thereby implying adequate convergence for each dimension (Hair, et, al 1995).

Table 3: Measure of Reliability

Item	Tangibles (1)	Reliability (2)	Assurance (3)	Empathy (4)
Cronbach's alpha	0.78	0.81	0.82	0.88

The relative influence of the 4 dimensions

The relative influence of the 4 dimensions on the overall service quality were investigated using the multiple regression analysis. Table 5 shows the result of the analysis where the dependent variable was overall service quality measured on a scale of 1 (very poor) to 7 (excellent). The results show that all four except one dimension, that is, reliability, are significant predictors of service quality.

Table 4: Construct Validity Indices

Criteria	Estimates
a) Criterion validity	Correlation with overall quality (.53) and overall satisfaction (.49) p <0.01
b) Discriminant validity	AVE ranges from 0.53 to 0.60 which are equal to or greater than square of correlation between two factors
c) Convergent validity	AVE ranges from 0.53 to .60 for the four dimensions. These AVEs are higher than .5

Table 5: Relative Importance of the 4 Dimensions in Predicting Service Quality

Dimension	Standardised coefficient (Beta)	P-value
Tangible	0.114	0.003 (significant)
Reliability	-0.041	0.294 (not significant)
Assurance	0.253	0.000 (significant)
Empathy	0.322	0.000 (significant)

Adjusted R-square= 29.8%

Summary and Conclusion

This paper reports on a study that attempts to develop an instrument, named ODLPERF to measure the performance of an ODL institution. Several tests such as normality, unidimensionality, reliability and validity tests were conducted to examine the appropriateness of ODLPERF as a performance measure for ODL institution. All the tests appear to indicate that it is an appropriate instrument that can be used to measure the performance of ODL institutions.

However, as mentioned at the outset of the paper, this is an initial exploratory attempt to develop the measuring instrument. As such, a number of rigorous tests to derive a full-blown measurement scale had not been attempted. In this regard, additional analyses need to be undertaken in order further refine this measuring instrument. In addition, the current study was conducted only in a single ODL institution, that is, OUM. For ODLPERF to have a general application, we need to extend the study to other ODL institutions.

References

Bolton, R.N. & Drew, J.H. (1991) A multi stage model of customer's assessments of service quality and value. *Journal of Consumer Research*, 17, 375–384.

Boulding, W., Kalra, A., Staelin, R. and Zeithaml, V.A. (1993), "A dynamic process model of service quality: from expectations to behavioural intentions", *Journal of Marketing Research*, 30, pp. 7-27.

Bryne, B.M. (1994) *Structural Equation Modelling with EQS and EQS/Windows-Basic Concepts, Applications and Programming*. Sage Publications, Thousand Oaks, CA.

Chow, J.C.C., Snowden, L.R. & McConnell, W. (2001) A confirmatory factor analysis of the BASIS-32 in racial and ethnic samples. *The Journal of Behavioural Health Services and Research*, 28, 400–411.

Comrey, A. L., & Lee, H.B. (1992) *A first course in factor analysis* (2nd edition) Hillsdale, NJ: Erlbaum

Cronin, J.J. & Taylor, S.A. (1992) Measuring service quality: reexamination and extension. *Journal of Marketing*, 56, 55–68.

Eisen, S.V., Wilcox, M. & Leff, H.S. (1999) Assessing behavioural health outcomes in outpatient programs: reliability and validity of the BASIS-32. *Journal of Behavioural Health Sciences and Research*, 26, 5–17.

Firdaus, A. (2005) The development of HEdPERF: A new measuring instrument of service quality for higher education. *International Journal of Consumer Studies*, 30 (6), 569-581.

Firdaus, A. (2006) Measuring service quality in higher education: HEdPERF versus SERVPERF. *Marketing Intelligence & Planning*, 24 (1), 31-47

Hair, J.F. Jr, Anderson, R.E., Tatham, R.L. & Black, W.C. (1995) *Multivariate Data Analysis with Readings*. Prentice Hall International Editions, Englewood Cliffs, NJ.

Kaiser, H.F. (1970) A second-generation little jiffy. *Psychometrika*, 35, 401–415.

Kelloway, E. (1998) *Using LISREL for Structural Equation Modeling: A Research Guide*, Sage, Thousand Oaks, CA.

Latifah, A.L., Ramli, B. & Ng Man San (2008a) "Sustainability of Post-Graduate Programmes at Open University Malaysia (OUM): An Importance-Performance Analysis." A paper presented at SEAAIR 2008 Conference, Surabaya, Indonesia.

Latifah, A.L., Ramli, B. & Ng Man San (2008b) "An Importance-Performance Survey (Postgraduate) Report." An unpublished report submitted to the Top Management of Open University Malaysia.

Lawrence S. Meyers, Glenn Gamst & A.J. Guarino (2006). Applied Multivariate Research: Design and Interpretation Sage Publication, Inc. Thousand Oaks, California

Llusar, J.C.B. and Zornoza, C.C. (2000), "Validity and reliability in perceived quality measurement models: an empirical investigation in Spanish ceramic companies", *International Journal of Quality & Reliability Management*, 17 (8), pp. 899-918.

Nunnally, J.C. (1988) *Psychometric Theory*. McGraw-Hill Book Company, Englewood-Cliffs, NJ.

Oliver, R.L. (1989), "Processing of the satisfaction response in consumption: a suggested framework and research propositions", *Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behaviour*, 2, pp. 1-16.

Parasuraman, A., Zeithaml, V.A. & Berry, L.L. (1988) SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64, 12–40

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1994), "Reassessment of expectations as a comparison standard in measuring service quality: Implications for future research", *Journal of Marketing*, 58, pp. 111-24.

Quester, P., Wilkinson, J.W. and Romaniuk, S. (1995), "A test of four service quality measurement scales: the case of the Australian advertising industry", Working Paper 39, Centre de Recherche et d'Etudes Appliquees, Group esc Nantes Atlantique, Graduate School of Management, Nantes.

Tabachnick, B.G. & Fidell, L.S. (1989) *Using Multivariate Statistics*, 2nd ed. New Harper-Collins Publishers, New York, NY.