

CONCERNS ABOUT SERVQUAL'S UNDERLYING DIMENSIONS

Mark A. Myerscough, Ph.D., Illinois State University, mark@themyerscoughs.com

ABSTRACT

The ServQual instrument has engendered considerable debate in MIS research articles. The ServQual instrument has been the target of repeated criticisms in its originating discipline (Marketing), as well as in a variety of other disciplines. This study focuses attention on the underlying constructs included in the ServQual instrument. The results of this study strongly suggest that an individual's expectations for service quality are expressed in a different set of factors than those contained in the same individual's perceptions of service performance. This fundamental difference between the expectations for service quality and the perceptions of service performance indicates that the calculation of the gap score-based ServQual index is inappropriate. Accordingly, this study recommends against the use of the ServQual gap score as the basis of information systems quality assessment.

Keywords: ServQual, Quality Assessment, Survey Validation, Gap Scores

INTRODUCTION

The ServQual instrument has been the target of considerable debate and utilization in Information Systems research recently, including multiple publications in MIS Quarterly ([17], [26], [27], [29]) and Decision Sciences ([16], [30]). These articles are in fact an extension of the long-running debate within the Marketing research discipline, which questions the validity and utility of the ServQual instrument ([3], [4], [6], [7], [9], [10], [12], [19]). This study is aimed at helping to explain the frequently conflicting results obtained with the ServQual instrument.

LITERATURE REVIEW

A consumer's perception of quality levels has long been a focus for marketing literature research. For example, the consumer's judgement concerning an entity's overall level of excellence or superiority has been used as a measurement of perceived quality [31]. Objective measures of quality, measured by elements such as the "conformance to requirements" [8] or "freedom from deficiencies" [15 p. 2.2] have been defined as the basis for quality assessment. However, these objective measures are difficult to translate into methods for assessing service (as opposed to product) quality. This difficulty led to the development of ServQual, intended to assess user perceptions of quality in a service environment [24].

Original Development of Survey Instrument (ServQual)

Parasuraman, Zeithaml and Berry [25] developed the original 22 item ServQual scale with questions intended to assess five specific dimensions (tangibles, reliability, responsiveness, assurance, and empathy). The ServQual instrument utilizes a "gap (or difference) score" analysis methodology, wherein the user's expectations for service quality are assessed at the same time as the user's perception of the actual system performance. The difference between these two scores (performance minus expectation) is used as the basis of analysis.

Multiple Marketing-oriented researchers ([2], [4], [9],[10], [19]) have identified factor stability as a problem for the ServQual instrument's assessment of service quality. At least two studies have found evidence that ServQual represents a unidimensional model ([6], [19]). A 1993 study concluded that the performance-only element of ServQual (referred to as ServPerf) "performs about as well as ServQual itself" ([3 p. 134]. The authors found that "Overall, the nomological validity evidence somewhat favors the non-difference score measure to the ServQual measure" [3 p. 136].

Survey Instrument's Adaptation to Information Systems

Following the initial development of the ServQual instrument, various researchers have attempted to adapt this instrument to their specific field of study. Pitt, Watson and Kavan argued that service quality "...needs to be considered as an additional measure of IS success." [26 p. 175-176] It is important to note that these authors did not state that service quality is the preferred method of assessing information systems quality, but rather an additional measure.

These results were disputed by Van Dyke, Prybutok and Kappelman ([29]), in a study identifying many of the deficiencies already cited ([3], [4], [6], [10]). The study found that the performance-only scores performed better than the performance minus expectation difference scores, consistent with the previous studies ([3]). Furthermore, the four dimensions of the ServQual instrument utilized in the study exhibited extremely poor fits when conducting exploratory factor analysis. The final conclusion of this study is that "the ServQual instrument, utilizing difference scores, is neither a reliable nor a valid measurement for operationalizing the service quality construct for an information systems services provider" [29 p. 204].

Table 1. Kettinger and Lee's [18] Criticism of the Van Dyke, Kappelman and Prybutok [29] Study

Criticism	Van Dyke, Kappelman and Prybutok's Implementation	This Study's Implementation
"Paucity of information regarding the administration of the survey"	Survey was administered prior to the principal researchers involvement	Survey was administered to a sample of employees.
The use of different item scales than the original 7 point Likert scale used by Kettinger and Lee (1995)	5 point Likert scale semantic differential	7 point Likert scale semantic differential
Use of external customers as opposed to internal IS users results in a different population group.	Survey administered to external IS customers	Survey administered to internal IS users.

Van Dyke, Prybutok and Kappelman exchanged research notes with Kettinger and Lee in 1999 ([18], [30]). Kettinger and Lee conclude their article by stating "We continue to take the position that well-established, managerially useful measures should not be discarded until such time as their underlying theory and practicality have been conceptually and empirically discredited" [18 p. 898]. This study provides additional justification of the position that the ServQual instrument is neither conceptually nor empirically valid for the assessment of information systems quality.

Kettinger and Lee identified several deficiencies in the Van Dyke, Kappelman and Prybutok study, all of which are addressed in this study. These issues are summarized in Table 1. Kettinger and Lee state

that “these limitations would appear to be so critical that they would defeat their own research purpose” [18, p. 895]. As summarized in Table 1, all of these issues are effectively dealt with in this study.

METHODS

This study focuses on a specific issue - the underlying, foundations for the gap score calculations comprising the ServQual instrument. The data was collected from employees in two Fortune 500 firms which were guaranteed anonymity during the course of data collection.

Participants

The survey instrument was distributed to internal information systems users of two Fortune 500 firms, and was focused on each organization's Web-Based Intranet System (W-BIS). Complete demographics are omitted due to space restraints, however the overall response rate exceeded 50%, with a total of 120 usable responses.

Apparatus

The ServQual index is a gap-score calculation based on a two-part survey instrument which requires the respondent to answer a matched set of questions such as:

Expectations: “Excellent W-BIS will maintain fully-functional systems and software”

Performance: “WBIS.XYZ.COM staff maintain fully-functional systems and software”.

The gap score is then calculated as (Performance response) minus (Expectations response). Since the ServPerf model utilizes this same measurement of Performance, it was expedient to include the analysis of ServPerf's stability in this analysis.

The original, revised five-factor ServQual model contained 22 items, segmented into five factors [25]. One factor (Tangibles) garnered consistently poor support in IS-based research projects ([27], [16]) and was dropped from this study. Therefore, only the 18 items comprising the original factors of reliability, assurance, responsiveness, and empathy were included.

Statistical Analysis

Three initial exploratory factor analyzes (ServQual - Expectations, ServQual - Performance, and ServQual gap scores) were conducted, allowing the extraction of only those factors with an eigenvalue greater than one. Hair, Anderson, Tatham and Black ([11]) recommend a process of first identifying the highest loading score for an individual item, and then determining whether this value exceeds the specified (.50 in this study) cutoff value.

RESULTS

Tables 2 through 4 summarize several significant values for each factor analysis. The Table's first three rows contain the initial eigenvalue for the factor, the eigenvalue after the varimax rotation, and the percentage of variance extracted for each rotated factor. These values are useful for determining how much of the variance included in the overall dataset was represented by the extracted factors. The Table's next four rows contain the four *a priori* factors (Reliability, Responsiveness, Assurance, and

Empathy) hypothesized in the ServQual instrument. For each *a priori* factor, the Table displays how individual items load on the extracted factors.

Structure of Tables Summarizing the Factor Analysis

The following description will help illustrate these tables. In Table 2, an examination of the *a priori* factor of Responsiveness illustrates that items six (Inform users of available services) and seven (Provide prompt services) both loaded on Factor A, while item nine (Never too busy to respond) loaded on Factor B. Item eight (Willing to assist users) did not show up as loading on either Factor A or Factor B.

The original factor loading scores for item 8 were .619 for Factor A and .613 for Factor B. These loading scores (both of which exceed the .50 cutoff value identified for this study) represent an item fairly evenly distributed between the two extracted factors. The final two rows of the table include how many different items (from the original eighteen) were identified as loading on this specific extracted factor, and the number of different *a priori* factors which were included in this extracted factor.

Convergent Validity Exhibited by Exploratory Factor Analysis

Maintaining only those eigenvectors for which the eigenvalue was greater than one as the extracted

Table 2. ServQual (Expectations Items) Exploratory Factor Analysis Summary

ServQual - Expectations		Extracted Factors		Number of Different Extracted Factors Represented
		A	B	
Eigenvalue prior to rotation		12.193	1.172	
Eigenvalue after Varimax		6.945	6.420	
% of variance after Varimax		38.582	35.668	
a priori Factors	Reliability (1-5)	1 – 5		1
	Responsiveness	6 – 7	9	2
	Assurance (10-13)		10 - 13	1
	Empathy (14-18)		14 - 18	1
Total Number of Items		7	10	
a priori Factors Included		2	3	

factor(s) in the exploratory factor analysis failed to identify the four factors hypothesized in ServQual model for any of the three sets of data analysis (ServQual - Expectations, ServQual - Performance, and ServQual). The ServQual - Expectations analysis identified two factors, as did the ServQual exploratory factor analysis. Analysis of the ServQual - Performance data identified three factors.

Table 2 summarizes the results of the ServQual - Expectation exploratory factor analysis. This table shows that three of the *a priori* factors were, in fact, extracted intact with all of the Reliability items loaded on Factor A and all of the Assurance and all of the Empathy items loaded on Factor B. However, the items associated with Responsiveness were split between both factors. Item eight (willing to assist users) was not shown included in this table since its factor loading scores were split nearly identically between the two extracted factors (.619 for Factor A and .613 for Factor B). These results indicate that while there appeared to be some internal convergent validity between the individual items that make up the Reliability, Assurance, and Empathy factors, there was a lack of discriminant validity between Assurance and Empathy.

In contrast to the results of the ServQual - Expectations analysis, the ServQual - Performance analysis (summarized in Table 3) was not neatly discernable. All four of the *a priori* factors had significant

factor loadings on at least two of the three extracted factors. This result indicated that the *a priori* factors showed an extremely low level of convergent validity. The poorest result was shown by the items which are contained in the Empathy factor, as individual items from this factor loaded significantly on all three of the extracted factors. As a result of the splitting of factors, the extracted factors had little discernable relationship with the hypothesized factors, with extracted Factor A including items from all four *a priori* factors, extracted Factor B including items from three of the *a priori* factors, and extracted Factor C containing items from two *a priori* factors.

Table 3. ServQual (Performance Items) Exploratory Factor Analysis Summary

ServQual – Performance		Extracted Factors			Number of Different Extracted Factors Represented
		A	B	C	
Eigenvalue prior to rotation		10.459	1.188	1.011	
Eigenvalue after rotation		5.487	4.568	2.604	
% of variance after rotation		30.481	25.37	14.465	
<i>a priori</i> Factors	Reliability (1-5)	1, 2, 4	3, 5		2
	Responsiveness	8, 9	6, 7		2
	Assurance (10-13)	10, 11, 12		13	2
	Empathy (14-18)	14, 16	18	15, 17	3
Total Number of Items		10	5	3	
<i>a priori</i> Factors Included		4	3	2	

Table 4. ServQual (Gap Scores) Exploratory Factor Analysis Summary

ServQual (Gap Scores)		Extracted Factors		Number of Different Extracted Factors Represented
		A	B	
Eigenvalue prior to rotation		11.464	1.078	
Eigenvalue after rotation		6.586	5.955	
% of variance after rotation		36.590	33.085	
<i>a priori</i> Factors	Reliability (1-5)		1, 2, 3, 4, 5	1
	Responsiveness (6-9)	8, 9	6, 7	2
	Assurance (10-13)	10, 11,		1
	Empathy (14-18)	14, 15,		1
Total Number of Items		10	7	
<i>a priori</i> Factors Included		3	2	

The exploratory factor analysis of the ServQual data is summarized in Table 4. The loading pattern for the ServQual data is very analogous to the loading pattern exhibited with the ServQual - Expectations data. The extracted factor loading patterns are nearly mirror-images of each other, with Factor B in Table 4 containing the exact same items as were contained in Factor A in the ServQual - Expectations

summary (Table 2). There are two differences in the factor labeled as Factor A in Table 4 and Factor B in Table 2. Item 8 (willing to assist users) now has a fairly definite factor loading, with a loading score of .726 for Factor A as opposed to a .437 for Factor B. However, Item 18 (Understands user's specific needs) was split identically (.536) between the two factors.

The same comment made for the ServQual - Expectations data holds for the ServQual data. The overall results indicate internal convergent validity between the items that make up the Reliability, Assurance, and Empathy factors, with a lack of discriminant validity between Assurance and Empathy. The factor Responsiveness exhibited a lack of convergent validity.

Summary of Convergent Validity Exhibited in Factor Analysis Results

The use of exploratory factor analysis was an attempt to identify the underlying factors contained within a set of data. In this study, it was utilized to identify the underlying factors contained in the ServQual and ServPerf data, and compare this result with the *a priori* - identified model. The results of the analysis showed that, in general, the four factor *a priori* model for ServQual and ServPerf were poorly fitted to the data collected in this study.

DISCUSSION

A variety of problems have been previously identified with the ServQual / ServPerf instruments. These problems were identified both in the original Marketing literature, in which the ServQual / ServPerf instruments were originally defined, and in the Information Systems literature.

As a result of all of the difficulties associated with the ServQual instrument, the present study conducted an analysis of the appropriateness of using this instrument in the assessment of Information Systems quality. Both the ServQual and the ServPerf instrument exhibited factor instability, with neither instrument generating a factor loading pattern consistent with the originally defined Service Quality model ([23]). Not only were the factor loading patterns generated in this study different from the hypothesized models, they were also inconsistent when compared with each other. The factor loading patterns for the expectations component of the ServQual instrument demonstrated a markedly different pattern than that of the ServQual performance component (see Tables 2 and 3). This result indicates that individuals utilized different internal models when identifying their expectations for service quality than when assessing the performance level of these same service attributes.

Van Dyke, Prybutok and Kappelman (1997) raised concerns about the psychological appropriateness of using the difference score methodology as the basis for simulating an individual's internal thought process. The factor scores computed in this study echo this concern. There is strong evidence to support the fact that users assess expectations for an information systems service level in a manner different from how the same user measures the actual performance quality of information systems services. Therefore, the calculation of the gap score difference between expectations and performance is an exercise in futility, since the underlying items contained in this calculation are not representative of a single, consistent factor. Since the assessment of actual performance levels did not fit the hypothesized model, the validity of calculating the gap scores based on this data is, at best, arguable. Additionally, the factor instability of the ServQual model argues strongly against the validity of averaging the individual elements gap scores to create the overall factor indices.

RECOMMENDATIONS

With all of the problems identified for the ServQual and ServPerf instrument, it is not possible for this study to recommend the continued utilization of this instrument as an acceptable research tool for the measurement of information systems quality. It is the recommendation of this study that the ServQual / ServPerf instrument not be used to assess the quality of information systems service levels, since the instrument has failed to exhibit acceptable levels of reliability and validity. Furthermore, the contradictory results for the evaluation of the ServQual expectations items in relation to the ServQual performance items raised serious questions about the appropriateness of difference-score measures as they relate to the assessment of expectations for and performance of quality-related attributes. Further research into the process used by individuals when assessing their expectations as opposed to actual performance is required.

REFERENCES

- [1] Babakus, Emen and Gregory W. Boller. (1992) "An Empirical Assessment of the ServQual Scale", *Journal of Business Research* (24). pp. 253-268.
- [2] Babukus, Emen, and W. Glynn Mangold. (1989) "Adapting the 'ServQual' Scale to Health Care Environment: An Empirical Assessment." in *AMA Educators' Proceedings* ed. Paul Bloom, et al., Chicago, IL: American Marketing Association.
- [3] Brown, Tom J., Gilbert A. Churchill, Jr., and J. Paul Peter. (1993) "Research Note: Improving the Measurement of Service Quality." *Journal of Retailing* (69.1) 127-139.
- [4] Carman, James M. (1990) "Consumer Perceptions of Service Quality: An Assessment of the ServQual Dimensions." *Journal of Retailing* (66.1) 33-55.
- [5] Cronbach, L. J. and P. E. Meehl. (1955) "Construct Validity in Psychological Tests." *Psychological Bulletin* (52.4) 281-302.
- [6] Cronin, J. Joseph Jr., and Steven A. Taylor. (1992) "Measuring Service Quality: A Reexamination and Extension", *Journal of Marketing* (56.7) 55-68.
- [7] Cronin, J. Joseph Jr. and Steven A. Taylor. (1994) "ServPerf versus ServQual: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality", *Journal of Marketing* (58.1) 125-131.
- [8] Crosby, Phillip B. (1979) *Quality is Free*. New York: New American Library.
- [9] Finn, David W. and Charles W. Lamb. (1991) "An Evaluation of the ServQual Scales in a Retailing Setting", in *Advances in Consumer Research*, Rebecca H. Holman and Michael R. Solomon, (eds.) (18).
- [10] Gagliano, Kathryn and Jan Hathcote. (1994) "Customer Expectations and Perceptions of Service Quality in Retail Apparel Speciality Stores", *Journal of Services Marketing* (8.1) 60-69.
- [11] Hair, Joseph F. Jr., Rolph E. Anderson, Ronald L. Tatham, and William C. Black. (1995) *Multivariate Data Analysis*. Englewood Cliffs, New Jersey: Prentice Hall.
- [12] Heung, Vincent C.S., M.Y. Wong, and Hailin Qu. (2000) "Airport-restaurant Service Quality in Hong Kong: An Application of SERVQUAL". *Cornell Hotel and Restaurant Administration Quarterly*. (41.3).
- [13] Johns, Gary. (1981) "Difference Score Measures of Organizational Behavior Variables: A Critique", *Organizational Behavior and Human Performance* (27.3) 443-463.
- [14] Juran, J. M. (Editor-in-chief) and Frank Gryna (Associate Editor). (1989) *Juran's Quality Control Handbook, 4th Edition*. United States of America: McGraw-Hill.
- [15] Juran, J. M. and Frank M. Gryna. (1993) *Quality Planning and Analysis, Third Edition*. New York: McGraw-Hill, Inc.
- [16] Kettinger, William J. and Choong C. Lee. (1995) "Perceived Service Quality and User Satisfaction with the Information Services Function." *Decision Sciences* (25.5-6) 737-766.
- [17] Kettinger, William J. and Choong C. Lee. (1997) "Pragmatic Perspectives on the Measurement of Information Systems Service Quality", *MIS Quarterly* (21.2) 223-240.
- [18] Kettinger, William J. and Choong C. Lee. (1999) "Replication of Measures in Information Systems Research: The Case of IS ServQual" *Decision Sciences* (30.3) 893-899.
- [19] Lam, Simon S. (1995) "Assessing the Validity of SERVQUAL: An Empirical Analysis in Hong Kong" *Asia Pacific Journal of Quality Management* (4.4). 33-40.
- [20] Myerscough, Mark A., Thomas Richards, and Jack D. Becker. (1997) "A Framework for Intranet-Based Information Systems (I-BIS) Applications." *Association for Information Systems - 1997 Americas Conference Proceedings*. Indianapolis, Indiana.
- [21] Parasuraman, A., Leonard Berry, and Valerie Zeithaml. (1991) "Refinement and Reassessment of the SERVQUAL Scale." *Journal of Retailing* (67.4) 420-450.
- [22] Parasuraman, A., Leonard Berry, and Valerie Zeithaml. (1993) "Research Note: More on Improving Service Quality Measurement", *Journal of Retailing* (69.1) 140-147.
- [23] Parasuraman, A., Valerie Zeithaml, and Leonard Berry. (1985) "A Conceptual Model of Service Quality and Its Implications for Future Research", *Journal of Marketing* (49.3) 41-50.
- [24] Parasuraman, A., Valerie Zeithaml, and Leonard Berry. (1985) "Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research" *Journal of Marketing* (58.1) 111-123.
- [25] Parasuraman, A., Valerie Zeithaml, and Leonard Berry. (1988) "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality." *Journal of Retailing* (64.1) 12-37.
- [26] Pitt, Leyland F., Richard T. Watson, and C. Bruce Kavan. (1997) "Measuring Information Systems Service Quality: Concerns for a Complete Canvas", *MIS Quarterly* (21.2) 209-222.
- [27] Pitt, Leyland F., Richard T. Watson, and C. Bruce Kavan. (1995) "Service Quality: A Measure of Information Systems Effectiveness", *MIS Quarterly* (19.2) 173-187
- [28] Van Dyke, Thomas and Michael Popelka. (1993) "Development of a Quality Measure for an Information Systems Provider", *Proceedings of the Decision Sciences Institute* (3) 1910-1912.
- [29] Van Dyke, Thomas P., Victor Prybutok, and Leon Kappelman. (1997) "Measuring Information Systems Service Quality: Concerns on the Use of the SERVQUAL Questionnaire", *MIS Quarterly* (21.2) 195 - 208.

- [30] Van Dyke, Thomas P., Victor Prybutok, and Leon Kappelman. (1999) "Cautions on the Use of the SERVQUAL Measure to Assess the Quality of Information Systems Services", *Decision Sciences* (30.3) 877 - 891.
- [31] Zeithaml, Valerie. (1987) Defining and Relating Price, Perceived Quality, and Perceived Value. Cambridge, Massachusetts: Marketing Science Institute.