December 2002

INFORMATION SYSTEMS QUALITY ASSESSMENT: REPLICATING KETTINGER AND LEES USISF/SERVQUAL COMBINATION

Mark Myerscough
Illinois State University

INFORMATION SYSTEMS QUALITY ASSESSMENT: REPLICATING KETTINGER AND LEE'S USISF/SERVQUAL COMBINATION

Mark A. Myerscough
Illinois State University
mark@themyerscoughs.com

Abstract

In a 1995 article, Kettinger and Lee proposed the combining of elements from two quality assessment instruments – the User Information Satisfaction instrument, and the ServQual (Service Quality) instrument into a single information systems quality assessment instrument. Through careful analysis and evaluation, the authors selected a sub-set of items from these two instruments in order to more fully assess the perceived quality of information systems than was possible with either of the individual instruments. This study reports on a replication of their original study, and analyzes their proposed combination. The results of confirmatory factor analysis suggests that this combined instrument is valid, and the study recommends further research into the utility of the combined (and abbreviated) UIS/ServQual instrument for the purposes of assessing information systems quality.

Introduction

In 1995 Kettinger and Lee published a study reporting on the use of a combined User Information Satisfaction (labeled as USISF – User Satisfaction with Information Systems Function in the paper) and a revised version of the Service Quality (ServQual) assessment instrument (originally proposed by Parasuraman, Zeithaml and Berry, 1985a) for the purposes of assessing information quality. Since that time, considerable research has been conducted and published concerning the suitability of the ServQual instrument for Information Systems quality assessment (Kettinger and Lee, 1997 & 1999; Pitt Watson and Kavan 1995 & 1997; Van Dyke, Prybutok, and Kappelman, 1997 & 1999; Van Dyke and Popelka, 1993). However, to the best of this researcher's knowledge, no one has ever conducted a rigorous assessment of Kettinger and Lee's original premise – which was that the combination of selected elements from Ives, Olsen and Baroudi's (1983) User Information Satisfaction instrument and from the ServQual instrument could be combined to assess the quality of an Information System. This study represents a confirmatory factor analysis evaluation of Kettinger and Lee's original contention – that the combination of elements from these two instruments is an appropriate quality assessment instrument.

Literature Review

The specific instrument evaluated in this study has been the subject of a single published research work – the article that reported on the original development. Therefore, the literature review for this paper covers three essential elements – the two original quality assessment instruments from which Kettinger and Lee derived their instrument, and the specific statistical techniques used in this study.

User Information Satisfaction

The construct of User Information Satisfaction (UIS) has been used as a surrogate for a variety of information systems quality measures in a large number of research projects since it was first developed in 1983. This utilization is consistent with its use in this study as an assessment of the user's perceptions of the overall quality of the Web-Based Information System (W-BIS). Melone (1990) attributes its application, in part, to two considerations. These are the provision of "a standard for making comparisons across organizational units and over time within units" and the fact that they are "relatively simple and inexpensive to administer" (Melone 1990, p. 76). Whether these were in fact the reasons for this acceptance and utilization, it is an indisputable fact that a large number of research projects in a variety of topics have incorporated the User Information Satisfaction construct as an appropriate system assessment metric.

The User Information Satisfaction (UIS) instrument is a seventeen-item questionnaire, which employs the use of semantic differential scales to assess the user's level of satisfaction with an information system. The instrument includes thirteen specific items, broken into three factors of Information Systems Personnel (five items), Information Product Quality (five items) and Knowledge and Involvement (three items). There are also three factor summary questions (one each for Information Systems Personnel, Information Product Quality, and Knowledge and Involvement) and a global satisfaction measure.

Klenke (1992) critiqued the construct validity for research into user satisfaction and user involvement. This critique focused attention on the process used by researchers in user satisfaction to select the component of user satisfaction. Klenke identified 18 different research projects, which were aimed at developing and/or validating an instrument for user satisfaction assessment. Of these 18 efforts, six used either the original Bailey and Pearson (1983) UIS instrument set of factors, or a modification of this scale designed to more closely fit the specific research project. Six other projects used either the Ives, Olson and Baroudi (1983) short-form instrument or an adaptation of this instrument designed to more closely fit the specific research project.

Since the Ives, Olson and Baroudi instrument is, itself, a derivative of the Bailey and Pearson instrument, this indicates that the original Bailey and Pearson work was responsible for, or a component in, almost every user information satisfaction scale designed since its initial publication in 1983. In fact, there were only two studies included in the Klenke analysis which were published following the development of these two instruments (Bailey and Pearson and Ives, Olson, and Baroudi) which did not rely on one of these two instruments for its construct measurement. One of these was the Doll & Torkzadeh (1988) end-user computing satisfaction instrument; the other is Rushinek & Rushinek's 1986 user satisfaction instrument. This illustrates the pervasive nature of the UIS instrument in terms of the assessment of information system quality.

Examples of using a measure of User Information Satisfaction for an overall quality surrogate in a research project include Barki and Huff's (1990) research into decision support systems usage patterns, Baronas and Louis' (1988) research into user involvement and systems acceptance, Beise's (1989) research concerning the assessment of information systems and organizational interfaces, Zviran's (1992) research into issues related to hospital information systems, and Benard and Satir's (1993) research concerning executive information systems and executive information needs. While this is not an exhaustive listing of UIS's utilization in research, it is indicative of its breadth of topics.

However, Galletta and Lederer (1989) have presented some reservations as to the applicability of this study as an overall systems effectiveness or efficiency measure. They identified three predominant reasons for the importance of user satisfaction. These were: the fulfillment of the MIS department's goals, the user's quality of life, and the extent of voluntary system utilization. All three of these components can be impacted upon by the user's level of satisfaction with an information system.

ServQual/ServPerf

A consumer's perception of quality levels has long been a focus for marketing literature research. For example, the consumer's judgment concerning an entity's overall level of excellence or superiority has been used as a measurement of perceived quality (Zeithaml, 1987). Objective measures of quality, measured by elements such as the "conformance to requirements" (Crosby, 1979) or "freedom from deficiencies" (Juran and Gryna 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 (Parasuraman, Zeithaml and Berry 1985a).

Original Development of Survey Instrument (ServQual)

Parasuraman, Zeithaml and Berry (1988) 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 (Babukus and Mangold, 1989; Carmen, 1990; Finn and Lamb, 1991; Gagliano and Hathcote, 1994; Lam, 1995) have identified factor stability as a problem for the ServQual instrument's assessment of service quality. Cronin and Taylor (1992) sought to investigate whether a single element of performance-based assessment, rather than the expectations minus performance difference was an appropriate instrumentation for measuring service quality. At least two studies have found evidence that ServQual represents a unidimensional model (Cronin and Taylor, 1992; Lam, 1995). A 1993 study concluded that the performance-only element of ServQual (referred to as ServPerf) "performs about as well as ServQual itself" (Brown, Churchill and Peter, 1993 p. 134). The authors found that "Overall, the nomological validity evidence somewhat favors the non-difference score measure to the ServQual measure" (Brown, Churchill and Peter, 1993 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. Relevant to this research project are those studies that have applied the ServQual instrument to the assessment of information systems quality. Pitt, Watson and Kavan argued that service quality "…needs to be considered as an additional measure of IS success." (Pitt, Watson and Kavan, 1997 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 serves as an additional measure.

These results were disputed by Van Dyke, Prybutok and Kappelman (1997), in a research study that identified many of the same deficiencies with the ServQual instrument already cited (Brown, Churchill, and Peter, 1993; Carmen, 1990; Cronin and Taylor, 1992; Gagliano and Hathcote, 1994). This study also included an analysis of ServPerf (the performance-only assessment) as an assessment for information systems service quality. The results of the analysis of ServPerf showed that the performance-only scores consistently performed better than the performance minus expectation difference scores, consistent with the results garnered by Brown, Churchill, and Peter (1993).

Table 1. Kettinger and Lee's (1999) Criticism of the Van Dyke, Kappelman and Prybutok (1999) 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.

A final problem identified in the Van Dyke, Prybutok and Kappelman study was that 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" (1997, p. 204).

Van Dyke, Prybutok and Kappelman exchanged research notes with Kettinger and Lee in a 1999 issue of Decision Sciences. 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" (1999, p. 898).

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, "these limitations would appear to be so critical that they would defeat their own research purpose" (1999, p. 895). As summarized in Table 1, all of these issues are effectively dealt with in this study.

Factor Analytic Techniques

Factor analytic techniques such as exploratory and confirmatory analysis are utilized in an attempt to reduce a large number of individual variables into a smaller set of constructs, or factors (Hair, Anderson, Tatham, and Black 1995). The factor analytic techniques were utilized in order to identify the underlying factors which are represented in these scales. In exploratory factor analysis, this is accomplished by calculating the factor loading scores and eigenvalues for the eigenvectors identified in the analysis. In confirmatory factor analysis, a pre-defined model is tested, and a number of statistics are evaluated which measure how well the data under analysis fits this pre-defined model. This study employed a confirmatory factor analysis of Kettinger and Lee's 1995 USISF/ServQual model.

In an investigation of the relationship between exploratory factor analysis and confirmatory factor analysis models, Gerbing and Hamilton used a simulation to generate test data known to exhibit specific factor loading patterns and sample sizes (Gerbing and Hamilton 1996). The test data was created so as to represent a robust test of the exploratory and confirmatory factor analysis techniques. The simulated data represented 100 replications of both 100 and 300 respondents for each of three different factor correlation patterns. In addition to the 100 replications of this 2 * 3 model (sample size * correlation patterns), an additional model was simulated which was based on a previously empirically validated model. This simulation used the factor loading values from the validated model to generate the simulated data and again generated 100 replications of the two sample sizes–100 and 300. Two different exploratory factor extraction techniques were used by Gerbing and Hamilton–principal components, since this is the most commonly utilized exploratory factor analysis technique, and maximum likelihood, since this most closely correlates with the maximum likelihood estimation commonly utilized in confirmatory factor analysis. In the final conclusion of this study was that "The most general conclusion of this article is that exploratory factor analysis is a useful tool to aid the researcher in recovering an underlying measurement model that can then be evaluated with confirmatory factor analysis" (Gerbing and Hamilton 1996, p. 71).

Methods

This study focuses on a specific issue - the verification of Kettinger and Lee's proposed model for information systems quality assessment (Kettinger and Lee, 1995). The data was collected from employees in two Fortune 500 firms. The companies were guaranteed anonymity during the course of data collection; therefore their names are not included in this report.

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 from this report due to space restraints, however the overall response rate for the surveys exceeded 50%, with a total of 120 usable responses. The instrument utilized a seven-point semantic differential scale, consistent with Kettinger and Lee's (1995, 1999) recommendation.

Apparatus

The survey instrument was approximately four pages long, and included a variety of questions concerning the individual's personal perceptions of information systems quality. Other components of the survey addressed issues such as the user's familiarity with information technology, their personal level of technology acceptance and confidence, and a variety of other personal factors. Of interest to this study was the inclusion of the complete User Information Satisfaction and ServQual instruments. Since Kettinger and Lee derived their entire model from a portion of the questions included in these instruments, all of the questions included in their model were included in this study. Additionally, since the entire instruments were utilized (consistent with Kettinger and Lee's survey), the study represents a complete replication of their previous research. Therefore, these results will allow for an independent analysis of the validity of Kettinger and Lee's USISF model.

The actual elements that made up the survey questions are detailed in Table 2. These items were based on Kettinger and Lee's original questions, however the wording was adjusted slightly to reflect the specific systems under observation in this study.

Table 2. Actual Survey Questions (ServQual questions reflect the performance-oriented question of the paired-items)

Factor Name	Variable Name	Actual Survey Question
Kn&Inv - UIS	UIS3	The degree of training provided to users by WBIS.XYZ.COM support staff
	UIS4	Your understanding of WBIS.XYZ.COM
	UIS5	Your feeling of participation in the functioning of WBIS.XYZ.COM
QIP-UIS	UIS7	The reliability of information, software, and documentation related to WBIS.XYZ.CO
	UIS8	Relevancy of information, software, and documentation related to WBIS.XYZ.COM
	UIS9	Accuracy of information, software, and documentation related to WBIS.XYZ.COM
	UIS10	Precision of information, software, and documentation related to WBIS.XYZ.COM
AESS-UIS	UIS6	The attitude of WBIS.XYZ.COM support staf
	UIS11	Communication with WBIS.XYZ.COM support staff.
Rel-SQ	SQ_1	When employees involved with WBIS.XYZ.COM promise to do something by a certain time, they do so
	SQ_3	WBIS.XYZ.COM staff perform requested services right the first time.
	SQ_4	WBIS.XYZ.COM staff provide their services at the time they promised to do so
Emp-SQ	SQ_14	WBIS.XYZ.COM services staff give you individual attention
	SQ_16	WBIS.XYZ.COM services has employees who give you personalized attention
	SQ_17	WBIS.XYZ.COM services staff have the your best interests at heart.
	SQ_18	The employees involved in WBIS.XYZ.COM service understand your specific needs

Statistical Procedures

This study was an attempt to replicate and confirm the findings of Kettinger and Lee (1995). Therefore, the focus of the analysis was the use of a series of Confirmatory Factor Analysis models. The statistical analysis was conducted with AMOS (release 4.01) software.

Models Analyzed with Structural Equation Modeling

In order to assess the existence of underlying factors of Kettinger & Lee's proposed USISF Model, a series of confirmatory factor analysis structural equation models were developed. These models, with their resulting analysis, are located in the Results section of this paper. The following discussion defines the basis for creating these models. To assist in reviewing the original definition of the model, all of the variable names, whenever possible, are identical to the original models from Kettinger and Lee's study (1995). The complete results generated from the analysis of the ServQual models are contained in the Results section of this paper. The actual survey questions are included in Table 2.

The first model (Model One - contained in Figure 1) illustrates all 16 of the selected items from the ServQual and User Information Satisfaction instruments loading on a single, global factor. This unidimensional result has been reported in previous research regarding the dimensionality of the ServQual instrument. This unidimensional pattern has occurred in both marketing (Cronin and Taylor 1992, 1994) and information systems (Van Dyke and Popelka 1993) research. Since one goal for this project was to conduct an analysis of how well the proposed model for information systems quality assessment performed when assessing the quality of a W-BIS, this unidimensional model was an appropriate beginning point. This single, global factor was named Overall Quality.

The second model used in this study illustrates the five a priori factors identified by Kettinger and Lee (made up of three factors from the original UIS instrument - Knowledge and Involvement, Quality of Information Product, and Support Service, along with two factors from the ServQual instrument - reliability and empathy). Model Two (Figure 2) shows these five factors as orthogonal elements of quality, not allowing for correlations between these factors. This lack of correlation is an unrealistic model, given that interactions between the elements that assessed quality would be expected, as depicted in Model Three. This model is included for the purposes of completeness in testing.

Model Three (contained in Figure 3) illustrates the existence of five correlated first-order factors in the proposed instrument. Allowing for non-directional correlations between these five factors allows for a level of association between the quality assessment factors.

Model Four (contained in Figure 4) includes the five first-order factors that all contribute to two higher level (second-order) constructs. This model represented a complete analysis of Kettinger and Lee's proposed model. Model Four shows that there are two interrelated quality issues, as assessed by the selected elements of the UIS and ServQual instruments.

The final model (Model Five) analyzed in this study is contained in Figure 5. Model Five illustrates that the two second-order factors (from Kettinger and Lee's model) were used as indicators of a global, overall quality construct. The rational for this model is that the two second-order factors (labeled as USISF and ServQual in Model 4) are components of an overall quality assessment for the information system in question.

"Fit" Indices and Measurements for Structural Equation Models

No single topic in the field of Structural Equation Modeling has generated as much attention as the issue of how to properly assess the validity of a structural equation model. Since the Confirmatory Factor Analysis form of a Structural Equation Model begins with an up-front, *a priori* specification of components and relationships, it is essential for the researcher to assess how well this a priori model "fits" the data undergoing analysis. The software package utilized for the structural equation modeling in this study (Amos 4.01 for Windows) calculates well over twenty different "Fit Indices" (as these assessment statistics are referred to), along with ninety percent confidence intervals for some of these indices.

It is also important to take into account Browne and Cudeck's cautionary note that, "Fit indices should not be regarded as measures of usefulness of a model. They contain some information about the lack of fit of a model, but none about plausibility" (Browne and Cudeck 1993, p. 157). In this way, the fit indices can be considered as analogous to measures of reliability. It is possible to have a highly reliable, and yet invalid measurement technique (for example an automobile odometer that is improperly calibrated - it provides the same results consistently, but these results are also consistently inaccurate). However, it is not possible to have a valid measurement technique that does not first exhibit a high level of reliability. A structural equation model that exhibits a good measurement of fit is not necessarily a valid model, but a structural equation model that does not possess good fit is, by definition, an invalid model.

Based on an analysis of relevant research, a total of seven indices were selected for reporting in this study to allow comparisons between the various structural equation models. Collectively, these Fit Indices and Fit Measures were used to assess the level of convergent validity for a given Structural Equation Model. A brief definition for each of these seven items is provided for the reader who is unaccustomed to or unfamiliar with Structural Equation Modeling.

Chi-square probability

The analysis of a Structural Equation Model focuses on the minimization of a discrepancy function, the structure of which requires a far more in-depth treatment than the space allotted here allows. In short, this discrepancy function (for maximum likelihood estimation) follows a Chi-Square (χ^2) distribution pattern and can be interpreted based on the degrees of freedom associated with the model. A probability value (P-value) for this statistic is "testing the hypothesis that the model fits perfectly in the population" (Arbuckle 1997, p. 554). While hypothesis testing is a well-known and widely accepted statistical technique, its unsuitability as a device for model selection was first identified in 1969 (Jöreskog 1969). Browne and Mels (1992) go so far as to offer the opinion that "this null hypothesis [of perfect fit] is implausible and that it does not help much to know whether or not the statistical test has been able to detect that it is false" (p. 78). Therefore, while this statistic was not relied on for model analysis and selection in this study, its value was reported for the sake of complete statistical reporting.

Root mean square error of approximation

The Root Mean Square Error of Approximation (RMSEA) incorporates an analysis of the complexity of a model into its calculation. This statistic is based on an estimate of the population discrepancy along with the number of degrees of freedom in the model. Browne and Cudeck (1993) opinioned that "a value of about 0.08 or less for the RMSEA would indicate a reasonable error of approximation" and that they "would not want to employ a model with a RMSEA greater than 0.1" (Browne and Cudeck 1993, p. 144). Based on this analysis, and desiring to provide the greatest latitude possible in the testing of the structural equation models, a maximum value of 0.1 was used as one of the criteria for identifying an acceptable fit for these models.

Minimum discrepancy/degrees of freedom ratio

For the maximum likelihood estimation technique, the ratio of χ^2/df should approach the value of one. However, as one author states, "The trouble is that it isn't clear how far from one you should let the ratio get before concluding that a model is unsatisfactory" (Arbuckle 1997, p. 555) Several rules of thumb are offered, with values ranging from a low-end estimate of two (Byrne 1989) to a high-end estimate of five (Wheaton, Muthen, Alwin, and Summers. 1977). This study employed a value of 3 for the χ^2/df ratio, which is consistent with other published Information Systems research efforts involving Structural Equation Models (Kettinger and Lee 1995; Segars and Grover 1993).

Root mean square residual

The Root Mean Square Residual (RMR) is the square root of the average squared amount by which the sample variances and covariances differ from their respective estimates that would be generated from a correct model. (Arbuckle and Wothke, 1999) An RMR value of zero indicates a model that has achieved "perfect" fit, and values greater than one are regarded as excessively high.

Goodness of fit index

The Goodness of Fit Index (GFI) was devised Jöreskog and Sörbom in 1984 and always ranges between zero and one. A value of one (unity) indicates a perfect fit, and a commonly accepted standard for this statistic is .90. A number of studies have shown that GFI is sensitive to sample sizes, leading to Anderson and Gerbing's conclusion that "Whereas GFI is independent of sample size (in that sample size is not an explicit part of the equation that defines GFI), the distribution of GFI values is strongly affected by sample sizes" (Anderson and Gerbing 1984, p. 172). This study found that as much as 28.3% of the variance in GFI values was attributable to variations in sample sizes, with the impact more pronounced with small sample sizes (n=50).

Comparative fit index

The Comparative Fit Index (CFI) was defined by Bentler (1990) and is based on a comparison of the hypothesized model to what is referred to as the "baseline" or "independent" model. This index can then be viewed as an analysis of how much "better" the hypothesized model is than the independent model. Values for CFI are forced to fall between the range of zero and one, as values exceeding these limits are truncated, and reported as zero or one, respectively.

Bentler (1990) reported the findings of a simulation/comparison study that utilized 200 replications of a structural equation model that investigated the impact of six different sample sizes ranging from 50 to 1600 in an exponential progression (50, 100, 200, etc.). This study found that CFI's standard deviation was consistently less than that of the other five indices, ranging from .034 at a sample size of 50, to .001 with a sample size of 1600. This result helped lead to the conclusion that:

> "CFI seems to be the best index; . . . it has a 0-1 range, has small sampling variability, and estimates the relative difference in noncentrality of interest. However, these advantages are obtained at the expense of some downward bias. This bias is quite small, and is certainly much less than the bias of the NFI (normed fit index). In fact, there was virtually no bias in the simulation with the misspecified model" (Bentler 1990, p. 245).

Based on this conclusion, CFI was selected as one of the leading fit indices for this study. CFI values less than 0.90 were considered to be strong indicators of a poorly fitted model.

Results

Since the purpose of this study is an attempt to replicate and confirm previous research, the results will be presented on the basis of the analysis of the five Models, previously defined. These models follow a logical "building-block" approach, beginning with a unidimensional model, which in essence tests whether there are, in fact, a variety of factors in the survey instrument. Acceptance of this model would indicate that the proposed five factor model does not possess strong factor discrimination. Based on the review of the results presented in Figure 1 (which are summarized in Table 3), the unidimensional model is rejected. All seven of the fit indices selected for inclusion in this study reject the acceptance of the unidimensional model. Of particular note is the RMSEA value in excess of .21, over double the stated limit of .10, along with the Comparative Fit Index value of .579, well below the stated minimum of .90.

The second model (Model 2 – displayed in Figure 2) is likewise rejected. This result was anticipated, since the Model propositions five independent factors, all related to information quality, without any interactions between these factors. However, the inclusion of the model does provide the second step in process of building an adequate model. While the fit indices were considerably improved over those in Model 1, they failed to achieve acceptable levels for any of the indices selected for this study. Based on this result, (improvement in all fit indices) there is evidence to suggest the existence of various factors in the data. Therefore, it was appropriate to continue the analysis with Models 3 through 5, all of which are built upon the five factor model identified by Kettinger and Lee (1995).

Model 3 is the first model for which the fit indices indicated an acceptable level of performance. Five of the seven fit indices included in this study identify this model as achieving an acceptable level of performance. The two indices that reject this model were the Chi-square probability test and the Goodness of Fit Index. The Chi-square probability test was already identified as "this null hypothesis [of perfect fit] is implausible and that it does not help much to know whether or not the statistical test has been able to detect that it is false" (Browne and Mel p. 78). Likewise, the Goodness of Fit Index, while popularly reported in a variety of structural equation-based articles, suffers from variations related to small sample sizes. The fact that both the Tucker-Lewis Index (.917) and the Comparative Fit Index (.935) values recommend acceptance of the model is important, since these indices were identified as among the stronger indices to use for this study.

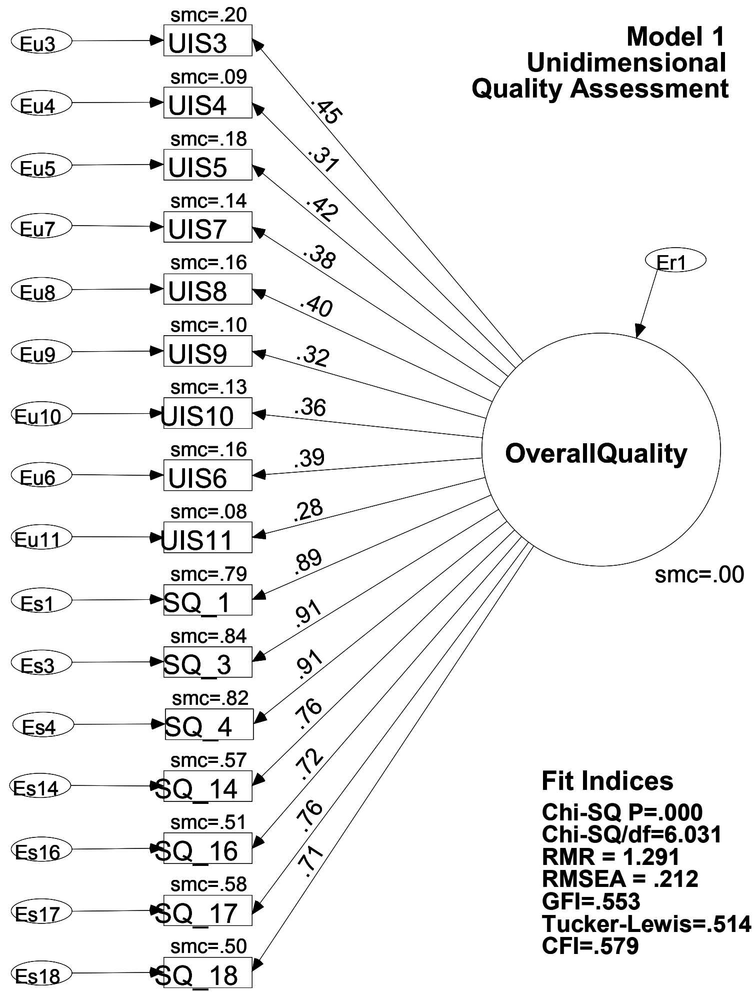

Figure 1. Unidimensional Model

Model 4 is also accepted based on the fit indices selected for this study, although at slightly less positive margins. While all five of the indices that registered an acceptable performance for Model 3 also showed an acceptable performance by Model 4, all of the values demonstrated slightly poorer results. However, none of the changes resulted in an index value that would represent a rejection of the model. However, consistent with the results for Model 3, the Chi-square probability test and the Goodness of Fit Index both failed to accept Model 4.

The fifth (and final) model for this study is an extension of Kettinger and Lee's original work. The Model tests whether the two separate second-order factors (labeled USISF and ServQual) in Model 4 are actually indicators of an over-arching construct (labeled as Overall Quality). Rather than testing for a correlation between the two factors (as Model 4 did), Model 5 tests for the existence of a overall quality dimension, which has been operationalized by the measurements contained in the USISF and ServQual instruments. This model represents an extension of the original work, and was not originally hypothesized by Kettinger and Lee. The results generated by this analysis are roughly equal to that of Model 4. The same five indices signaled nearly acceptable levels of performance.

Table 3 summarizes the results for all five models. This table shows that for Models 3, 4, and 5 the same five goodness of fit indices signaled an acceptable level of performance, and summarizes the performance for all seven of the selected goodness of fit indices.

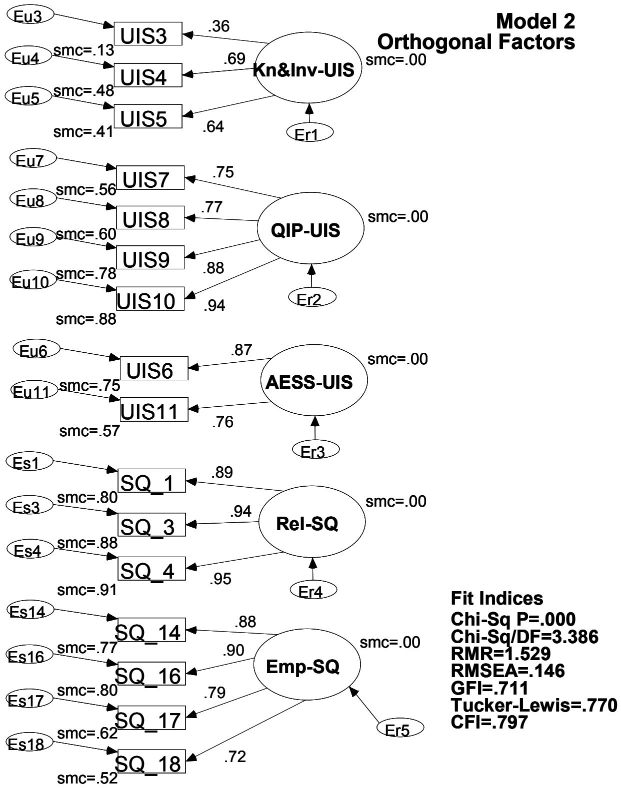

Figure 2. Orthogonal Constructs

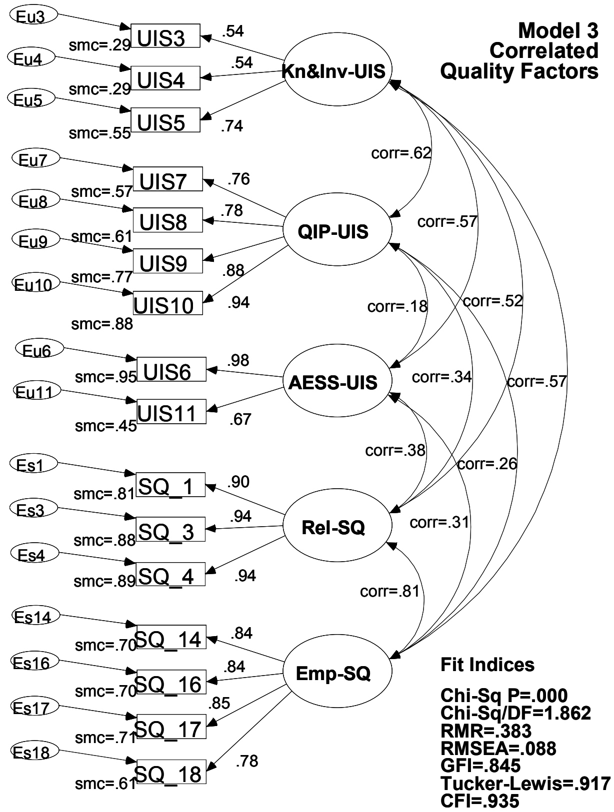

Figure 3. Correlated First-Order Factors

Discussion

The results of the Confirmatory Factor Analysis gives support to Kettinger and Lee's acceptance of elements of the ServQual instrument as an additional element for assessing the quality of Information Systems. Based on the results of this study, the use of a 27 item survey for the assessment of information systems quality is recommended. This survey would include the nine individual items from the original UIS instrument, fourteen items from ServQual instrument (comprised of seven items each for expectations and performance) and the four global quality assessment questions traditionally included in the UIS instrument. This questionnaire, while longer than the original 17 item UIS instrument, is considerably shorter than the combination of the complete UIS and ServQual instrument.

Additional research is needed to test the robustness of this survey, as well as the results generated when administering the abbreviated format since both Kettinger and Lee's original study (1995) and this study administered the full UIS instrument combined with the complete ServQual instrument. It is potentially possible that the inclusion of these additional questions influenced the responses of the individuals involved in this study. Therefore, the next step in this research will be the administration of the limited questionnaire.

The results for Models 4 and 5 identify a potential problem with this model, in that Model 3 achieved better results. This result suggests that rather than assessing a larger scale construct such as overall quality, that these questions merely assess the existence of five separate, yet interrelated constructs with considerable interaction. It will require further research with the instrument to identify the specifics related to this result. However, the fact that the fit indices were still positive (i.e. recommending acceptance) for Models 4 and 5 indicates that while Model 3 is potentially a better explanation for the relationships observed in the data, that Models 4 and 5 are also acceptable explanations.

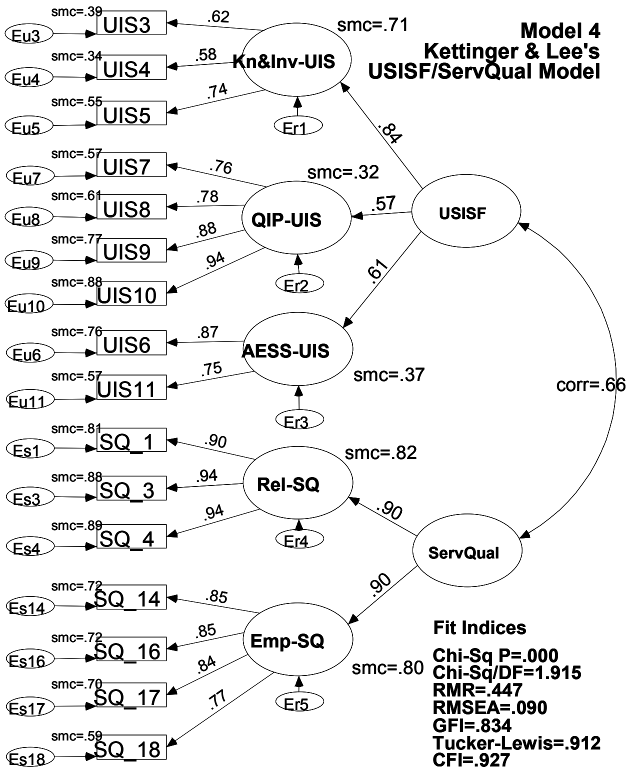

Figure 4. Correlated Second-Order Factors

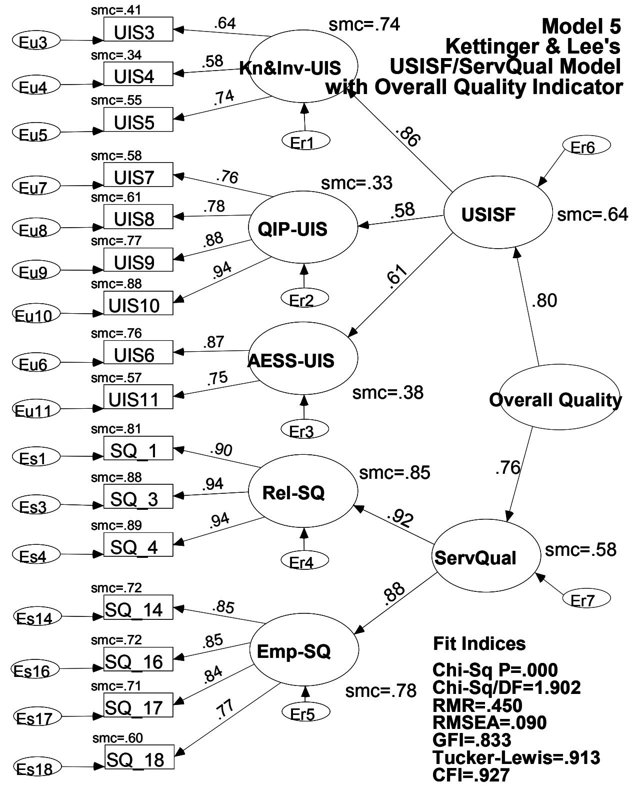

Figure 5. Overall Quality Indicator

This study has focused entirely on the assessment of the measurement model, based on the analysis of the fit indices generated in the analysis. The next step for the analysis will be to assess the actual structural model, to identify the magnitude and significance of the relationships identified in this study. This analysis is ongoing, and will be reported on in the presentation of this paper at the conference in August.

Table 3. Summary of Goodness-of-Fit Indices for Combined UIS/ServQual Models

Evaluation Statistic	Assessment Guideline	Model 1	Model 2	Model 3	Model 4	Model 5	Number of Accepts
χ^2's P-value	p > = .05	0.000	0.00	0.00	0.00	0.00	0
χ^2 / d.f.	< = 3	6.031	3.386	**<u>1.862</u>**	**<u>1.915</u>**	**<u>1.902</u>**	3
Root Mean Square Residual (RMR)	< = 1	1.291	1.529	**<u>.383</u>**	**<u>.447</u>**	**<u>.450</u>**	3
Root Mean Square Error of Approximation (RMSEA)	< = .1	.212	.146	**<u>.088</u>**	**<u>.090</u>**	**<u>.090</u>**	3
Goodness of Fit Index (GFI)	>= .9	.553	.711	.845	.834	.833	0
Tucker-Lewis Index (TLI)	>= .9	.514	.770	**<u>.917</u>**	**<u>.912</u>**	**<u>.913</u>**	3
Comparative Fit Index (CFI)	>= .9	.579	.797	**<u>.935</u>**	**<u>.927</u>**	**<u>.927</u>**	3
Number of indices that achieved or exceeded assessment guidelines		0	0	5	5	5	
Number of indices that failed to achieve assessment guidelines		7	7	2	2	2	

<u>Bold, underlined</u> items represent a fit measure which signals an acceptable performance.

References

Anderson, James C. and David W. Gerbing. 1984. The Effect of Sampling Error on Convergence, Improper Solutions, and Goodness-of-Fit Indices for Maximum Likelihood Confirmatory Factor Analysis. Psychometrika 49 155-173.

Anderson, James C. and David W. Gerbing. 1988. Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. Psychological Bulletin 103.3 411-423.

Arbuckle, James L. 1997. Amos Users' Guide, Version 3.6, Chicago, Illinois: Smallwaters Corporation.

Arbuckle, James L. 1999. Amos Users' Guide, Version 4.0, Chicago, Illinois: Smallwaters Corporation.

Babakus, Emen and Gregory W. Boller. 1992. An Empirical Assessment of the ServQual Scale. Journal of Business Research 24 253-268.

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.

Barki, H. and S. L. Huff. 1990. Implementing Decision Support Systems: Correlates of User Satisfaction and System Usage. INFOR 28.2 (May): 89-100.

Baronas, Ann-Marie K. and Meryl Reis Louis. 1988. Restoring a Sense of Control During Implementation: How User Involvement Leads to System Acceptance. MIS Quarterly 12.1 (March): 111-124.

Benard, Roch and Ahmet Satir. 1993. User Satisfaction with EISs: Meeting the Needs of Executive Users. Information Systems Management 10.4 (Fall): 21-29.

Bentler, P. M. 1990. Comparative Fit Indexes in Structural Models. Psychological Bulletin 107.2 (February): 238-246.

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 (Spring): 127-139.

Browne, Michael W. and G. Mels. 1992. RAMONA Users Guide. Columbus, Ohio: Ohio State University.

Browne, Michael W. and Robert Cudeck. 1993. Alternative Ways of Assessing Model Fit. in Testing Structural Equation Models, Kenneth A. Bollen and Scott Long, editors. Newbury Park, California: Sage Publications.

Carman, James M. 1990. Consumer Perceptions of Service Quality: An Assessment of the ServQual Dimensions. Journal of Retailing 66.1 (Spring): 33-55.

Cronin, J. Joseph Jr., and Steven A. Taylor. 1992. Measuring Service Quality: A Reexamination and Extension. Journal of Marketing 56.7 (July): 55-68.

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 (1994): 125-131.

Crosby, Phillip B. 1979. Quality is Free. New York: New American Library.

Doll, William J. and Gholamreza Torkzadeh. 1988. The Measurement of End-User Computing Satisfaction. MIS Quarterly 12.2 (June): 259-274.

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.

Galletta, D. F. and Albert Lederer. 1989. Some Cautions on the Measurement of User Information Satisfaction. Decision Sciences 20.3 (Summer): 419-439.

Gerbing, David W. and James C. Anderson. 1993. Monte Carlo Evaluations of Goodness-of-Fit Indices for Structural Equation Models. in Testing Structural Equation Models, Kenneth A. Bollen and Scott Long, editors. Newbury Park, California: Sage Publications.

Gerbing, David W. and Janet G. Hamilton. 1996. Viability of Exploratory Factor Analysis as a Precursor to Confirmatory Factor Analysis. Structural Equation Modeling 3.1 (Spring): 62-72.

Hair, Joseph F. Jr., Rolph E. Anderson, Ronald L. Tatham, and William C. Black. 1995. Multivariate Data Analysis. Englewood Cliffs, New Jersey: Prentice Hall.

Ives, Blake, Margrethe H. Olson, and Jack J. Baroudi. 1983. The Measurement of User Information Satisfaction. Communications of the ACM 26.10 (October): 785-793.

Jöreskog, Karl G. 1969. A General Approach to Confirmatory Maximum Likelihood Factor Analysis with Addendum. Psychometrika 34 183-202.

Jöreskog, Karl G. and Dag Sörbom. 1984. LISREL-IV User's Reference Guide. Mooresville, Indiana: Scientific Software.

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.

Juran, J. M. and Frank M. Gryna. 1993. Quality Planning and Analysis, Third Edition. New York: McGraw-Hill, Inc.

Kettinger, William J. and Choong C. Lee. 1995. Perceived Service Quality and User Satisfaction with the Information Services Function. Decision Sciences 25.5-6 (Sept:Oct-Nov:Dec): 737-766.

Kettinger, William J. and Choong C. Lee. 1997. Pragmatic Perspectives on the Measurement of Information Systems Service Quality. MIS Quarterly 21.2 (June): 223-240.

Kettinger, William J. and Choong C. Lee. 1999. "Replication of Measures in Information Systems Research: The Case of IS ServQual" Decision Sciences 30.3 (Summer): pp. 893-899.
Klenke, Karin. 1992. Construct Measurement in Management Information Systems: A Review and Critique of User Satisfaction and User Involvement Instruments. INFOR. 30.4 (November): 325-348.
Melone, Nancy P. 1990. A Theoretical Assessment of the User-Satisfaction Construct in Information Systems Research. Management Science. 36.1 (January): 76-91.
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. August 1997.
Parasuraman, A., Leonard Berry, and Valerie Zeithaml. 1991. Refinement and Reassessment of the SERVQUAL Scale. Journal of Retailing 67.4 (Winter): 420-450.
Parasuraman, A., Leonard Berry, and Valerie Zeithaml. 1993. Research Note: More on Improving Service Quality Measurement. Journal of Retailing 69.1 (Spring): 140-147.
Parasuraman, A., Valerie Zeithaml, and Leonard Berry. 1985a. A Conceptual Model of Service Quality and Its Implications for Future Research. Journal of Marketing 49.3 (Fall): 41-50.
Parasuraman, A., Valerie Zeithaml, and Leonard Berry. 1985b. Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research. Journal of Marketing 58.1 (January): 111-123.
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 (Spring): 12-37.
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 (June): 209-222.
Pitt, Leyland F., Richard T. Watson, and C. Bruce Kavan. 1995. Service Quality: A Measure of Information Systems Effectiveness. MIS Quarterly 19.2 (June): 173-187
Segars, Albert H. and Varun Grover. 1993. Re-Examining Perceived Ease of Use and Usefulness: A Confirmatory Factor Analysis. MIS Quarterly 17.4 (December): 517 - 525.
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.
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 (June): 195 - 208.
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 (Summer): pp. 877 - 891.
Zeithaml, Valerie. 1987. Defining and Relating Price, Perceived Quality, and Perceived Value. Cambridge, Massachusetts: Marketing Science Institute.
Zviran, Moshe. 1992. Evaluating User Satisfaction in a Hospital Environment: An Exploratory Study. Health Care Management Review 17.3 (Summer): 51-62.