

# Can't See the Trees for the Forest? Why IS-ServQual Items Matter

Ahmad A. Rabaa'i<sup>a,\*</sup>, Mary Tate<sup>b</sup>, Guy Gable<sup>c</sup>

<sup>a</sup> Assistant Professor, Department of Computer Science and Information Systems, College of Arts and Sciences, American University of Kuwait (AUK), Safat, Kuwait

<sup>b</sup> Senior Lecturer, School of Information Management Victoria University of Wellington, New Zealand

<sup>c</sup> Professor, School of Information Systems Queensland University of Technology (QUT), Brisbane, Australia

---

## ABSTRACT

Despite longstanding concern with the dimensionality of the service quality construct as measured by ServQual and IS-ServQual instruments, variations on the IS-ServQual instrument have been enduringly prominent in both academic research and practice in the field of IS. We explain the continuing popularity of the instrument based on the salience of the item set for predicting overall customer satisfaction, suggesting that the preoccupation with the dimensions has been a distraction. The implicit mutual exclusivity of the items suggests a more appropriate conceptualization of IS-ServQual as a formative index. This conceptualization resolves the paradox in IS-ServQual research, that of how an instrument with such well-known and well-documented weaknesses continue to be very influential and widely used by academics and practitioners. A formative conceptualization acknowledges and addresses the criticisms of IS-ServQual, while simultaneously explaining its enduring salience by focusing on the items rather than the "dimensions." By employing an opportunistic sample and adopting the most recent IS-ServQual instrument published in a leading IS journal (virtually, any valid IS-ServQual sample in combination with a previously tested instrument variant would suffice for study purposes), we demonstrate that when re-specified as both first-order and second-order formatives, IS-ServQual has good model quality metrics and high predictive power on customer satisfaction. We conclude that this formative specification has higher practical use and is more defensible theoretically.

*Keywords:* IS-ServQual, Formative Models, Customer Satisfaction, Service Quality, IS Evaluation

---

## I. Introduction

The IS-ServQual instrument in its various variants contains a number of inherent paradoxes that have puzzled researchers for three decades and which have

been debated in multiple studies in leading information systems journals (e.g., Kettinger and Lee, 1995; Kettinger and Lee, 1997; Klein et al., 2009; Pitt et al., 1995). In particular, the instrument has been criticized for (among other things) unstable

---

\*Corresponding Author. E-mail: [arabaai@auk.edu.kw](mailto:arabaai@auk.edu.kw). Tel: 96522248399

dimensionality (e.g., Babakus and Boller, 1992; Carman, 1990; Gorla et al., 2010; Jia and Reich, 2012; Jia et al., 2008; Jiang et al., 2002; Klein et al., 2009; Parasurman et al., 1991; Sylvester et al., 2011; Teas, 1994; Van Dyke et al., 1997; Yang and Peterson, 2004;). Yet leading variants have collectively been cited thousands of times, this being the paradox of IS-ServQual research; it is perhaps the most enduringly popular and influential stream of “flawed” research in the information systems discipline.

The ServQual instrument originated in the 1980s, in the then new research field of services marketing. ServQual is a multi-item scale developed to assess customer perceptions of service quality in service businesses (Parasuraman et al., 1988). Service quality was claimed to consist of five constructs (often described as “RATER”): reliability, the ability to perform service dependably and accurately; assurance, the ability of staff to inspire confidence and trust; tangibles, the physical facilities, equipment and staff appearance; empathy, the extent to which caring, individualized service is given; and responsiveness, the willingness to help and respond to customer needs.

The ServQual instrument “crossed the disciplinary barrier” from services marketing to information systems when it was applied by Pitt et al., (1995) as IS-ServQual to measure the service quality of the information systems function (ISF). The rise of e-commerce, and the increased role of information systems and self-service in delivering services that were previously delivered face to face, resulted in a resurgence of interest in service quality measurement (Sylvester et al., 2011) and the appropriation (or as some would argue mis-appropriation) of ServQual-related measures by a whole new generation of e-commerce systems researchers. ServQual-related studies continue to influence the provenance of recent work reported in leading in-

formation systems journals (e.g., Jia et al., 2012; Klein et al., 2009; Yang et al., 2011).

However, concern with the validity of ServQual has been as enduring as its popularity. In particular, the dimensionality of ServQual and the stability of the five RATER constructs across contexts and studies, has been hotly debated in both marketing and information systems literature (e.g., Carman, 1990; Jiang et al., 2002; Kettinger and Lee, 1997; Van Dyke et al., 1997) with various arguments proposed for alternative combinations of the items within the RATER or other dimension configurations.

As researchers, we are encouraged to generate and test theories at higher levels of abstraction (“the forest”) beyond individual, concrete and particular phenomena (“the trees”). This tends to result in privileging theories that purport to discover “constructs” or “dimensions” that are generalizable and invariant across multiple contexts. The preoccupation of researchers with the RATER dimensions is an example of this.

Herein, we argue that the instability and probable miss-specification of the RATER dimensions, is a distraction from the main contribution of the instrument. We argue that the enduring popularity and managerial diagnostic capability for information systems managers offered by IS-ServQual, derives primarily from the coverage provided by the census of *indicators* included in the IS-ServQual instrument. We further argue that the implicit independence of the items suggests more appropriate conceptualisation of IS-ServQual as a formative index; the inherent context-specificity of formative items offering a plausible explanation for mixed results observed across studies. In attention to these propositions, our research questions are:

1. Is the overall set of IS-ServQual indicators effective

in predicting customer satisfaction with the Information Systems Function (ISF<sup>1</sup>), when specified as a formative index?

2. Does each of the individual items make a unique contribution to the predictive power of IS-ServQual?
3. What (if any) is the role of the sub-dimensions (themes) of IS-ServQual?

### 1.1. Service Quality Dimensions

The dimensional instability of ServQual has been widely discussed by both marketing and information systems researchers and is summarised here briefly. Carman (1990) criticised ServQual for lack of dimensional stability when the instrument was applied in new settings. Critics suggested that the five dimensions did not have adequate discriminant validity, and were concerned that the combined results of ServQual studies would be so divergent that no clear theory would emerge (Babakus and Boller, 1992; Carman, 1990; Teas, 1994). In response to critics, the ServQual authors refined and continued the development of ServQual (Parasuraman et al., 1991). The 1991 revision resulted in three key claims: the ongoing diagnostic value of ServQual; an improved instrument addressing concerns about negatively worded statements; and reaffirmation of the validity of the original RATER factors (Parasuraman et al., 1991). Nevertheless, challenging the stability of the RATER dimensions became a regular point of departure for many ServQual based studies. Babakus and Boller (1992) found only two distinct dimensions, while Carman (1990) found eight. This argument was re-litigated by information systems researchers (Van

Dyke et al., 1997). An interesting summary of this and some of the other key “stories” emerging from three decades of ServQual-related studies is provided by Sylvester et al. (2011).

The issues with discriminant validity and dimensional instability are now widely acknowledged, even in “positive” citations of ServQual (e.g., Jiang et al., 2002), perhaps at the cost of scholarly attention to other theoretical issues. In recent studies in leading IS journals, issues with dimensionality continue. Amongst studies that revalidated the dimensions (many did not), support was found for the popular (in information systems) four dimension conceptualisation by Pitt et al. (1995) which excludes tangibles (e.g., Gorla et al., 2010; Jia and Reich, 2012; Jia et al., 2008; Klein et al., 2009), as well as a three dimension conceptualisation (excluding responsiveness) by Yang et al. (2011).

ServQual measures customer attitudes and perceptions towards service quality. Since these attitudes and perceptions are internal psychological states, they cannot be observed and measured directly. Psychometric literature distinguishes two types of relationships between indicators and latent variables - formative and reflective specification (Jarvis et al., 2003). Observations of reflective indicators are caused by the latent variable; if the value of the underlying latent variable changes, the values of the reflective indicators will also change. All reflective indicators of a construct are supposed to be “reflective” of the same concept, and thus interchangeable (Jarvis et al., 2003). Formative specification of indicators instead assumes that the observations of indicators collectively ‘form’ the variable<sup>2</sup>, which is an index of

1) Though we have with this study chosen to make the ISF the focal phenomenon of study, this is not consequential. We might readily have studied some other service - e.g., e-commerce.

2) There is theoretical debate about the ontology of formative variables. Some authors claim that formative variables are not “latent variables” (variables that are assumed to be real, but cannot be observed directly) at all, but should be considered

its indicators (Diamontopoulos and Winklhofer, 2001). Formative indicators can be independent of one another, and it is sensible to talk about them covering the 'scope' of a phenomenon of interest.

When the ServQual authors talk of constructs, they are by implication referring to latent variables, as they are referring to unobservable perceptions and attitudes held by customers. Tate and Evermann (2010)'s conceptual paper examines the five "RATER" constructs (latent variables) of ServQual and their indicators, and argues that the constructs appear to be a mix of formative and reflective indicators. If the RATER dimensions were truly defensible as valid and distinct latent variables, the dimensionality would be much more stable, and the face validity would be better - each dimension would have a clear, uni-dimensional definition, and be described by a set of interchangeable reflective indicators that were clearly different in meaning to the indicators of each of the other latent constructs (Borsboom et al., 2003).

We note that the original provenance of the ServQual items was a large-scale qualitative study including executive interviews and focus groups which was aimed at covering the range and scope of perceptions of service quality (Parasuraman et al., 1985). That study identified ten determinants. The RATER dimensions were not hypothesized a priori, but were derived, some might say reified, from exploratory factor analysis and scale purification. The definitions for the RATER dimensions contain "and" clauses, which suggest potentially composite rather than uni-dimensional constructs. The authors noted that the two "dimensions" of assurance and

empathy contain "*items representing seven original dimensions - communication, credibility, security, competence, courtesy, understanding/knowing customers, and access*" (Parasuraman et al., 1988, p. 23).

Although a reflective specification of the dimensions is the most popular, empirical support has been found for other conceptualizations of service quality with different psychometric properties. The four-dimension, fourteen-item scale proposed by Kettinger and Lee (1997) uses an overall "service" construct, which is composed of the four sub-dimensions of reliability, assurance, empathy and responsiveness, with no additional indicators at a global level. The four sub-dimensions are claimed to have convergent and discriminant validity, although subsequent confirmatory factor analysis did not produce good fit metrics (Kettinger and Lee, 1997). While not explicitly stated as such, "service" in that study can be assumed to be a second order formative index (Mackenzie et al., 2011) of four reflective sub-dimensions. Specified in this way, "service quality" is a composite construct which is a mathematical "construction" or "index" of its sub-dimensions (Borsboom, 2005; Bagozzi, 2011; MacKenzie et al., 2011).

In summary, there is a considerable body of evidence supporting the need to re-specify IS-ServQual. This evidence includes 1) continuing popularity of the instrument suggests that the indicators have strong diagnostic relevance for management, and the instrument appears to provide a complete "census" of relevant indicators in the manner of formative constructs, 2) instability of the constructs when specified reflectively suggests that reflective specification might be incorrect, 3) empirical support for service quality as a second order formative construct, formed by its first order dimensions (Kettinger and Lee, 1997), and 4) theoretical support for specifying the

---

as indexes or mathematical composites (Borsboom, 2005). We largely concur. The problem with this conception is that any and all indexes are equally theoretically valid. We argue that in this context the formative RATER "variables" are best reconceptualised as themes which link a set of related but independent competencies

dimensions wholly or partially as formative (e.g., Tate and Evermann, 2010).

## 1.2 The SERVQUAL Instrument and Information Systems Service Quality

IS researchers have used SERVQUAL instrument for different purposes. In fact, studies in IS altered the SERVQUAL measure to the specific context under study, either through modification of the indicators, deletion of dimensions, or addition of dimensions. For example, some have used the SERVQUAL instrument to demonstrate an interpretation of formative measurement in information systems research (Cenfetelli and Bassellier, 2009), while others (e.g., O'Cass and Carlson, 2012) have used the same instrument to provide an empirical assessment of consumers' evaluations of web site service quality. Also, Po-An Hsieh et al. (2012) studied the impact of user satisfaction with mandated CRM use on employee service quality.

The use of SERVQUAL instrument in IS research started when Kettinger and Lee (1995) established a short form (13 items) of SERVQUAL within the IS Function (ISF) context, that evidenced the validity of the five SERVQUAL dimensions. They, however, reserved claims of external validity, given the context specificity of their single sample.

Pitt et al. (1995) independently analysed SERVQUAL data across three different sample sites using principal components and maximum likelihood methods, thereby deriving alternative three-, five- and seven-factor solutions. Given their findings, Pitt et al. (1995, p. 181) report that "SERVQUAL does not clearly delineate among the dimensions of service quality." They warn users of the 22-item SERVQUAL to be aware of the co-alignment of the dimensions of responsiveness, assurance, and empathy due to their semantic

similarity, and observe that the reliability of the tangibles dimension is low. More broadly, the use of IS SERVQUAL has been a subject of considerable debate (e.g., Fisk, et al., 1993; Kettinger and Lee, 1995; Parasuraman et al., 1993; Pitt et al., 1997; Van Dyke et al., 1997; VanDyke et al., 1999). A focus of the debate has concerned calculating differences between two possibly different constructs, expectations and perceptions.

Kettinger and Lee (1997) conducted an empirical comparison between SERVQUAL and SERVPERF (SERVQUAL using only perceived scores - no gap measure) in terms of their relative psychometric superiority in the IS setting. While slightly better reliability and explained variance were noticed with the SERVPERF measures, neither SERVQUAL nor SERVPERF data have explained the tended fit of the SERVQUAL five factor structure. Nonetheless, SERVQUAL has been touted for its practical relevance (Jiang et al., 2002; Kettinger and Lee, 2005) and continues to be used to evaluate technical support service interactions (Carr, 2002). IS researchers have also expanded the use of service quality into new areas, such as measuring service quality longitudinally (Watson et al., 1998) and internationally (Kettinger et al., 1995). More recently, IS researchers have also adapted service quality for use in the evaluation of electronic service environments, such as e-commerce web sites (Li et al., 2003; Wang and Tang, 2003) and Internet banking sites (Jayawardhena, 2004).

Kettinger and Lee (2005) reported on a study of an alternative instrument adapted from marketing referred to as the "zones of tolerance" (ZOT) service quality measure. The authors argued that this zones of tolerance measure is conceptualised to overcome a central criticism of the original SERVQUAL instrument; namely, the need for a more parsimonious

<Table 1> The Four Dimensions of IS ZOT SERVQUAL (Kettinger and Lee, 2005)

Dimension	Explanation
Reliability	The ability to perform promised ISF services dependably and accurately
Responsiveness	The willingness to help ISF users and to provide prompt service
Rapport	The ISF ability to convey a rapport of knowledgeable, caring, and courteous support
Tangibles	Physical facilities, equipment, and appearance of personnel

conceptualisation of service quality expectations, while retaining the practical diagnostic power from gauging service expectation levels. By using a factor analysis technique, four constructs with 18 items were derived. Three original SERVQUAL constructs emerged from the exploratory factor analysis (tangibles, reliability, and responsiveness). However, two of the original dimensions, empathy and assurance, were merged into a fourth dimension. Kettinger and Lee (2005, p. 612) argued that, based on a review of the retained items and the seeming similarity of the constructs when applied in the IS context, *“the new merged construct was named rapport because the construct items focus on an IS service provider’s ability to convey a rapport of knowledgeable, caring, and courteous support.”* IS ZOT SERVQUAL contains measures for desired, adequate, and perceived service quality levels, and includes 18 items.

The IS ZOT SERVQUAL instrument was pretested through a series of interviews with IS professionals and IS graduate students. After pre-testing and refining the instrument, two samples were chosen for cross validation, an initial sample from the university setting and a holdout sample from the industry setting. The findings represent an important step toward addressing past concerns with the original IS SERVQUAL’s expectation measure and gap-scoring. The IS ZOT SERVQUAL instrument has strong practical potential as a diagnostic tool through

which managers can quickly visualize their current IS service quality situation and design corrective actions.

### 1.3. Information Systems Service Quality and Customer Satisfaction

Perhaps the strongest explanation for the enduring popularity of the IS-ServQual measures is their efficacy in predicting customer satisfaction. Customer satisfaction is defined in marketing research as the consumer’s fulfilment response (Oliver and Swan, 1989). In information systems, satisfaction has been of interest mainly with regards specific information systems (IS), but also with regards the information systems function (ISF). It is noted that affect towards an IS or ISF is a generalised response that provides little insight into antecedents of that response. The dimensions and items of service quality are much more granular, with greater diagnostic value. With the goal of assessing both the predictive power and nomological validity of IS service quality (i.e., identification through structural relations)<sup>3),4)</sup>, ‘satisfaction’

3) A nomological network includes a theoretical framework of research objects, an empirical framework of how these objects will be measured, and specification of the relationships between these two frameworks (Campbell and Fiske, 1959).

4) Nomological validity is evidenced where the formative construct behaves within a net of hypotheses as expected (Diamantopoulos and Winklhofer, 2001; Henseler et al.

is included in the study model as its immediate consequence. Following is a summary of the main IS satisfaction measures and instruments.

Doll and Torkzadeh (1988) define user satisfaction with an IS as the affective attitude towards a specific IS application by someone who interacts with the application directly. User satisfaction is probably the most extensively used single measure for IS evaluation (e.g., Au et al., 2002; Au et al., 2008; Briggs et al., 2008; Doll and Torkzadeh, 1988; Etezadi-Amoli and Farhoomand, 1996; Gatian, 1994; Igbaria and Nachman, 1990; Igbaria and Tan, 1997; Iivari, 1987; Ives et al., 1983; Sedera and Tan, 2005). However, despite the large amount of research that has been done on IS user satisfaction, with several widely cited studies and standard instruments that measure user satisfaction with IS (e.g., Bailey and Pearson, 1983; Baroudi and Orlikowski, 1988; Doll and Torkzadeh, 1988), several authors (e.g., Au et al., 2002; Goodhue, 1995; Iivari, 1997; Khalifa and Liu, 2004; Sedera and Tan, 2005; Woodroof and Kasper, 1998) note a range of conceptual problems related to IS user satisfaction instruments, and empirical evidence of their validity has yielded mixed and contradictory results (Delone, 1988; Klenke, 1992; Mahmood and Becker, 1985). Zviran and Erlich (2003) claim that the IS user satisfaction concept is used to refer both to the IS function and to a single IS application without always making clear the distinction. Zviran and Erlich (2003, p. 87) suggest that “*measures of user satisfaction with the information system function suffer from severe limitations as a measure of user satisfaction with a single application.*”

In this study, we conceptualise satisfaction as an immediate consequence of the service quality of the ISF (IS Service Quality) with the dual goals of assess-

ing IS Service Quality's nomological validity<sup>5)</sup> and its predictive power. The notion of satisfaction as immediate consequence of IS service quality has support in the marketing literature. Gable et al. (2008, p. 388) state that: “*Services marketing researchers (e.g., Anderson and Sullivan, 1993; Brady et al., 2005; Spreng and Mackoy, 1996) employ a nomological net that positions Satisfaction as an immediate consequence of Service Quality; Satisfaction being antecedent of Behavioural Intention.*”

In adoption research, it has been noted that theories of user attitudes and perceptions towards technology concentrate too narrowly on measures of beliefs and affects, and are “*leading to research that is unable to provide actionable advice*” (Benbasat and Barki, 2007, p. 213). A construct for user satisfaction with IS does not inform those involved in the management of IS, as to what characteristics or competencies will lead to satisfaction. One explanation for the popularity of IS-ServQual is that it meets this need. We argue that the individual IS-ServQual *items* represent the real contribution of the IS-ServQual instrument, and the opportunity for resolving the ServQual paradox. Thus, IS-ServQual-type measures can be seen to make a valuable contribution to the forgotten “left hand side” of models of attitudes towards information systems (Evermann and Tate, 2010), thereby providing insights into the detailed, descriptive beliefs that are antecedent to more generalised attitudes (Tate and Evermann, 2010). The original ServQual *items* were initially derived from extensive qualitative research (Parasuraman et al., 1985) and have been enduringly popular for their diagnostic capabilities. The original ServQual *dimensions*, by contrast, were likely reified from exploratory factor

---

2009; Urbach and Ahlemann, 2010).

---

5) Further discussion on IS service quality nomological validity is presented in Section 5.

analysis<sup>6),7)</sup> and have proved unstable (e.g., Carman, 1990; Hughey et al., 2003; Mels et al., 1997). As a case in point, the IS-ServQual instrument selected for this study (Kettinger and Lee, 2005), is an example of a study that finds a different factor structure to that originally proposed by the ServQual authors, despite almost three decades of ServQual research.

## II. Model Development

### 2.1. Specifying Formative vs. Reflective

Heightened awareness of the difference between formative and reflective constructs has cast doubt on the validity of many mainstream constructs employed in IS research over the past three decades (Petter et al., 2007). There is a significant threat of misspecifying and validating constructs as “reflective” that on closer scrutiny are in fact “formative (Gable and Sedera, 2009). Misspecification of constructs as formative or reflective results in measurement error, which impacts the structural model, thereby increasing the potential for type I and type II errors (Diamantopoulos and Winklhofer, 2001; Jarvis et al., 2003; Petter et al., 2007).

Reflective constructs have observed measures that are affected by an underlying latent, unobservable construct, while formative constructs are a composite of multiple measures (MacCallum and Browne, 1993; Petter et al., 2007). Bollen and Lennox (1991) demon-

strate the formative construct as:

$$Y = \beta_1 X_1 + \dots + \beta_n X_n + \zeta$$

Where: Y = the formative construct being estimated

$\beta_i$  = beta weight for indicators

$X_i$  = indicator scores/observations

$\zeta$  = a disturbance term

One main difference between reflective and formative indicators is the extent to which an indicator contributes to the construct under investigation (Jarvis et al., 2003; Petter et al., 2007). Hence, formative indicators are assigned beta weights (Petter et al., 2007), as shown in the equation above. Consequently, “*dropping a measure from a formative-indicator model may omit a unique part of the conceptual domain and change the meaning of the variable, because the construct is a composite of all the indicators*” (MacKenzie et al., 2005, p. 712). That is removing a nonsignificant formative indicator will remove the beta weight associated with it, no matter how large or small it might be (Petter et al., 2007, p. 627). Conversely, formative items are not expected to covary significantly (though they may, due to ‘common cause’), as there is an expectation that each will make a unique contribution to understanding the phenomenon of interest.

### 2.2. Multidimensional Constructs

Multidimensional constructs are characterised by involving more than one dimension (Edwards, 2001b; Jarvis et al., 2003; Law and Wong, 1999; Law et al., 1998; Netemeyer et al., 2003; Petter et al., 2007; Polites et al., 2011; MacKenzie et al., 2005; Wetzels et al., 2009), as opposed to unidimensional constructs which have a single underlying dimension (Netemeyer

---

6) The process of factor identification and scale purification followed by Parasuraman et al. (1988) appears vulnerable to this criticism.

7) It is noted that reasonably strong patterns of co-variation among the IS-ServQual items is typically observed (though these patterns vary across studies, thus concern with dimensionality). Though formative items need not co-vary, co-variance among formative items is often observed, typically due to ‘common cause’ (Petter et al., 2007).



et al., 2003; Polites et al., 2011; Wetzels et al., 2009). Edwards (2001b, p. 144) refers to multidimensional constructs as “*several distinct but related dimensions treated as a single theoretical concept,*” where each “*dimension represents a unique content domain of the broader construct*” (Polites et al., 2011, p. 1). However, the use of multidimensional constructs has generated considerable debate in the literature. Advocates of multidimensional constructs claim they “*provide holistic representations of complex phenomena, allow researchers to match broad predictors with broad outcomes, and increase explained variance*” (Edwards, 2001a, p. 145). Additionally, multidimensional constructs allow for more theoretical parsimony (Edwards, 2001b; Law et al., 1998; MacKenzie et al., 2005; Wetzels et al., 2009) and allow matching the level of abstraction for predictor and criterion variables (Edwards, 2001b; Wetzels et al., 2009). Critics of multidimensional constructs suggest they are “*conceptually ambiguous, explain less variance than explained by their dimensions taken collectively, and confound relationships between their dimensions and other constructs*” (Edwards, 2001a, p. 145). Regardless of this debate, Polites et al. (2011, p. 2) argue that multidimensional constructs “*... provide opportunities to advance IS research by enabling the capture of complex concepts in comparatively simple abstractions... Due to their potential to advance theory, multidimensional constructs have appeared with more frequency in top IS journals in recent years.*”

As mentioned previously, constructs are described as multidimensional when their indicators are themselves constructs or dimensions (Edwards, 2001b; Jarvis et al., 2003; Law and Wong, 1999; Petter et al., 2007; Polites et al., 2011). A basic distinction between the types of multidimensional constructs is the direction of the relationship between the construct and its dimensions (e.g., Edwards, 2001b; Law

and Wong, 1999; Petter et al., 2007; Polites et al., 2011). If the relationships point from the construct to its dimensions, the construct is referred to as *superordinate* because “*it represents a general concept that is manifested by its dimensions*” (Edwards, 2001b, p. 145), and each dimension “*represents a different manifestation or realisation of the underlying construct*” (Polites et al., 2011, p. 5). On the other hand, if the relationships point from the dimensions to the construct, the construct is referred to as *aggregate* because “*it combines or aggregates specific dimensions into a general concept*” (Edwards, 2001b, p. 145). Put differently, dimensions of an aggregate construct are combined to form the construct (Law et al., 1998), and these dimensions are analogous to formative measures (Edwards, 2001b)<sup>8</sup>.

### 2.3. Conceptualising IS Service Quality in this Study

The way a multidimensional construct is operationalized may influence analytical results from model testing (e.g., Gable and Sedera, 2009; Jarvis et al., 2003; Petter et al., 2007; Polites et al., 2011; Vlachos and Theotokis, 2009). Therefore, it is essential to carefully conceptualise the relationship between the first-order dimensions and their indicators and between lower-order dimensions and the higher-order construct (Polites et al., 2011). In this study, the IS service quality construct is conceptualised as

8) While a multidimensional construct (superordinate or aggregate) is conceptualised based on its dimensions “*it does not mean that it can exist separately from them*” (Edwards, 2001b, p. 145). The relationships between a multidimensional construct and its dimensions “*represent associations between a general concept and the dimensions that represent or constitute the construct,*” rather than “*causal forces linking separate conceptual entities*” (Edwards, 2001b, p. 146).

a multidimensional construct; both first-order formative and second-order formative. We use as the starting point, the shorter, four dimension, eighteen item instrument adapted from Kettinger and Lee (2005) included as <Table 2>.

This instrument was selected because it: 1) is the most recent of major IS-ServQual revisions published in a leading IS journal; and 2) is a version of IS-ServQual that preserves the original focus on the quality of the IS function. Though the K&L model

deviates from the original RATER dimensions by combining assurance and empathy into “rapport,” this is inconsequential for our purposes herein, which will be made clear. IS service quality is conceptualised as first-order formative, with the four Kettinger and Lee (2005) first-order dimensions (i.e., reliability, responsiveness, rapport and tangibles) having formative indicators. We do so on the basis that the dimensions manifest demonstrable instability when specified reflectively (see earlier), and in light of con-

<Table 2> The IS Service Quality Construct (Kettinger and Lee, 2005)

Dimension	Explanation
Reliability	The ability to perform promised ISF services dependably and accurately
Indictor 1	Providing services as promised
Indictor 2	Dependability in handling user's service problems
Indictor 3	Performing service right the first time
Indictor 4	Providing services at the promised time
Indictor 5	Maintaining reliable technology and system
Indictor 6	Prompt service to users
Responsiveness	The willingness to help ISF users and to provide prompt service
Indictor 1	Willingness to help users
Indictor 2	Readiness to respond to user's requests
Rapport	The ISF ability to convey a rapport of knowledgeable, caring, and courteous support
Indictor 1	Making users feel safer in computer transactions
Indictor 2	IS employees who are consistently courteous
Indictor 3	IS employees who have the knowledge to answer users' questions
Indictor 4	Giving users individual attention
Indictor 5	IS employees who deal with users in a caring fashion
Indictor 6	Having the user's best interest at heart
Indictor 7	IS employees who understand the needs of users
Tangibles	Physical facilities, equipment, and appearance of personnel
Indictor 1	Visually appealing facilities
Indictor 2	IS employees who appear professional
Indictor 3	Useful support materials (such as documentation, training, videos, etc...)

cerns with face validity of the dimension item combinations as reflective (Tate and Evermann, 2010). In example of the latter, with reference to the reliability dimension in <Table 2>, the ISF may provide services at the promised time (indicator 4), but it may not be maintaining reliable technology and system (indicator 5); thus the items need not co-vary and 'on the face of it' do not measure the same thing. In further example, in the tangibles dimension, the ISF may have visually appealing facilities (indicator 1), but it may not provide useful support materials (indicator 3). Thus, reliability, responsiveness, support and tangibles are here conceptualised as formative first-order dimensions.

At this point, we need to ask whether the first-order formative dimensions add any value to the model? We have argued for the importance and salience of the *items* relative to the dimensions. What purpose is served by first order formative dimensions? Ideally, all exogenous items in a formative model should be allowed to freely correlate with all others. However, there are both theoretical and practical arguments for including first-order formative "dimensions." From a theoretical perspective, having a large number of uncorrelated indicators is not very parsimonious. In our model, these dimensions provide thematic focus for their sets of indicators. Indicators that can, and do, vary independently, may yet share a common theme. Our first-order formative dimensions can be best conceptualised as containers for a basket of independent, but thematically related characteristics. From a practical perspective, in the same way that many reflectively specified indicators are imperfectly correlated with model quality heuristics providing "acceptable" levels of convergent and discriminant validity, it may be unreasonable to assume that formatively specified indicators are all perfectly uncorrelated, especially for a broad phenomenon such

as service quality. Models with large numbers of uncorrelated indicators may be difficult to identify (Petter et al., 2007), and the chances of some of the indicators making a non-significant contribution to the dependent variable increases. There is therefore value in modelling first-order formative variables, so long as we remember they are themes, or indexes, not real "dimensions" in the reflective sense. We hereafter refer to the first order formative constructs as "themes" to distinguish our approach from the reflective dimensions that have been the subject of previous research controversy.

While we have herein conformed to the dimensions/indicators employed by Kettinger and Lee (2005), this is not overly consequential for our central argument. Alternative combinations of items might also yield reasonable formative model metrics. These groupings are pragmatic, for interpretation, communication and subsequent action. In example, a reasonable rationale for the grouping of a particular set of items might be that they are jointly addressable by a common intervention. We consider the theme of reliability, which includes measures for "performing the service right first time" and "performing the service at the promised time." A thought experiment will show that these measures are independent and not interchangeable (it is possible to do the right thing at the wrong time, or vice versa, for example a major software upgrade in the first week of the semester). Nevertheless, both measures could likely be improved by better staff training and more robust processes (e.g., obtain sign-off from key stakeholders on functionality and timing before scheduling an upgrade).

We therefore model the IS service quality construct as formed from four, formative first-order themes (it is second-order formative). Aggregating these four, assumes that the themes collectively contribute to

the second-order IS service quality construct. However, these four themes are likely to change over time and be affected in different ways by different antecedents. For instance, IS service quality may be at different levels of effectiveness and efficiency. As such, one would be mistaken to easily trade, for example, reliability of the ISF for the responsiveness of the ISF. Further, a change in the tangible aspects of the ISF, for example, does not imply a similar change in the ISF ability to convey a rapport of knowledgeable, caring, and courteous support, thereby again making a reflective specification less compelling. In other words, the four themes that form the IS service quality construct are not interchangeable. Hence, this implies that reliability, responsiveness, rapport and tangibles affect IS service quality in a formative way. Accordingly, the IS service quality construct is conceptualised as a formative-second order construct.

In this study, the satisfaction construct is measured using four indicators adopted from the overall satisfaction scale developed by Spreng et al. (1996) based in Expectation-confirmation theory (ECT) (Oliver, 1981), considered a central theory for explaining satisfaction in marketing research (Cenfetelli et al., 2008).

### III. Research Method

#### 3.1. Instrument Design

The IS-ServQual instrument developed by Kettinger and Lee (2005) was used as the basis for this study. Seven-point Likert scales from 1 (strongly disagree) to 7 (strongly agree) were used to evaluate respondents' perceptions of the IS service quality (perceptions only)<sup>9</sup>.

The Satisfaction construct adopted, based on

Spreng et al.'s (1996) dimensions, has been validated in the IS context (e.g., Bhattacharjee, 2001; Bhattacharjee and Premkumar, 2004; Cenfetelli et al., 2008; Premkumar and Bhattacharjee, 2008). This construct captures respondents' satisfaction levels both in intensity and direction (Oliver, 1993; Oliver, 1997) along 7-point scales anchored between four semantic differential adjective pairs: "*frustrated/contente*," "*displeased/pleased*," "*terrible/delighted*," and "*dissatisfied/satisfied*" (Bhattacharjee, 2001).

#### 3.2. Data Collection

Data was collected from students of the Faculty of Information Technology at two well established public universities in Jordan<sup>10</sup>. Student subjects were used because they constitute a homogeneous group from an 'occupational stage of lifecycle' viewpoint. They also have an ongoing relationship with the central university ISF. They contact the ISF for a range of services, such as connecting to the university-wide network, enrolment support, support of computer laboratories, consulting, training, and normal help-desk assistance.

An anonymous, self-administered, hardcopy survey instrument was distributed to 265 undergraduate IT students in the two Jordanian universities; 219 usable responses were collected (82.6% response rate)<sup>11</sup>.

---

9) Kettinger and Lee (2005, p. 614) state that "*in cases where brevity, cost, or predictive validity concerns demand, the seemingly less clinical perception-only [...] measure might be a better option*," as is done in this study.

10) The data collection in Jordan involved translation effort, as the country's official language is Arabic. Several translation techniques are reported in the literature (e.g., Brislin, 1970; Brislin, 1986; Hansen, 1987; Samaddar and Kadiyala, 2006). Brislin's (1986) translation technique was thought to be the most appropriate for the context of this study.

The survey was divided into four sections. In the first section, respondents were asked several demographic questions for classification purposes. In the second section, respondents were asked to indicate the frequency of their contact with their ISF and if they were familiar with the ISF services<sup>12)</sup>. The third section measured perception of the ISF's service

quality. A 7-point Likert scale was used to elicit responses to each of the 18 items in <Table 1>, from 1 (strongly disagree) to 7 (strongly agree). The fourth section measured respondents' satisfaction with the ISF and included the four semantic differential adjective pairs (Bhattacharjee, 2001). The descriptive statistics are shown in <Table 3>.

<Table 3> Descriptive Statistics

IS Service Quality Items		Mean	Standard Deviation
Reliability Indicators			
Indicator 1	Providing services as promised	2.98	2.02
Indicator 2	Dependability in handling user's service problems	3.51	2.33
Indicator 3	Performing service right the first time	2.14	1.95
Indicator 4	Providing services at the promised time	2.68	1.79
Indicator 5	Maintaining reliable technology and system	3.25	2.07
Indicator 6	Prompt service to users	2.21	1.01
Responsiveness Indicators			
Indicator 1	Willingness to help users	3.45	1.55
Indicator 2	Readiness to respond to user's requests	3.51	1.65
Rapport Indicators			
Indicator 1	Making users feel safer in computer transactions	3.45	1.93
Indicator 2	IS employees who are consistently courteous	2.98	1.13
Indicator 3	IS employees who have the knowledge to answer users' questions	2.05	1.29
Indicator 4	Giving users individual attention	3.69	2.01
Indicator 5	IS employees who deal with users in a caring fashion	3.16	1.98
Indicator 6	Having the user's best interest at heart	2.36	1.48
Indicator 7	IS employees who understand the needs of users	3.54	2.55
Tangibles Indicators			
Indicator 1	Visually appealing facilities	3.45	2.47
Indicator 2	IS employees who appear professional	3.21	2.16
Indicator 3	Useful support materials (such as documentation, training, videos, etc.)	1.55	1.02
Satisfaction Indicators			
Indicator 1	Frustrated/contented	3.58	1.70
Indicator 2	Displeased/pleased	4.15	1.78
Indicator 3	Terrible/delighted	4.04	1.65
Indicator 4	dissatisfied/satisfied	3.95	1.69

## IV. Results

To assess the validity of IS service quality as a multi-dimensional formative construct, partial least squares (PLS) was conducted using SmartPLS 2.0 M3 (Ringle et al., 2005). PLS was chosen in this study because of its ability to readily model both formative<sup>13)</sup> and multidimensional constructs (Wetzels et al., 2009).

Whether or how formative constructs should be empirically and statistically validated is contended (Diamantopoulos et al., 2008, p. 13). Some researchers state that no quantitative, quality tests are applicable for evaluating the appropriateness of formative indices. Other researchers note that the appropriateness and applicability of statistical procedures is limited because the choice of formative indicators determines the conceptual meaning of the construct. For instance, Rossiter (2002, p. 315) questions the need for any validity assessment of formative indicators, claiming “*all that is needed is a set of distinct components as decided by expert judgment.*”

We and other researchers (e.g., Diamantopoulos et al., 2008; Edwards and Bagozzi, 2000; Götz et al., 2010; Henseler et al., 2009; Jarvis et al., 2003; MacKenzie et al., 2005; Petter et al., 2007; Rabaa'i and Gable, 2012) do not share this view, and stress the need to assess the validity of formative models.

However, traditional reflective construct validity assessments do not apply to formative models (e.g., Albers, 2010; Ali et al., 2012; Diamantopoulos, 1999, 2006; Diamantopoulos and Siguaaw, 2006; Götz et al., 2010; Rabaa'i and Gable, 2012).

Diamantopoulos (2006, p. 11) states, with respect to formative models, “*reliability becomes an irrelevant criterion for assessing measurement quality.*” It is the assumption of error-free measures that makes the question of indicator reliability irrelevant (Henseler et al., 2009). Unlike reflective indicators, indicators in a formative structure have no measurement error but rather a disturbance term, which represents the remainder of the construct domain unexplained by the presented indicators (Andreev et al., 2009, p. 5).

While reliability becomes an irrelevant criterion for assessing formative models (e.g., Bollen, 1984; Bollen, 1989; Diamantopoulos, 2006; Diamantopoulos and Siguaaw, 2006), the examination of validity becomes essential (Diamantopoulos, 2006; Diamantopoulos and Siguaaw, 2006; Götz et al., 2010; Henseler et al., 2009; Urbach and Ahlemann, 2010). Previous studies (e.g., Andreev et al., 2009; Götz et al., 2010; Henseler et al., 2009; Petter et al., 2007; Urbach and Ahlemann, 2010) suggest that the assessment of formative measurement models should entail (1) assessment at the indicators level and (2) assessment at the construct level. Thus, assessment of the IS service quality construct is conducted in three stages: assessment of the first-order formative indicators; assessment of the second-order formative themes; and assessment of IS service quality at the construct level.

### 4.1. Assessment of First-Order ‘Formative’ IS Service Quality

Assessing formative constructs raises the concern of whether each indicator contributes to the formative

11) Students were told that participation in the study was voluntary. However, the study was promoted by two senior authorities in these two universities, which accounts for the high response rate.

12) Only respondents who were familiar with various ISF services and who have frequently used these services were considered in the data analysis of this study. 14 collected surveys were excluded due to lack of such familiarity.

13) It should be noted that LISREL can handle formative models (Jarvis et al., 2003); however, such models are often easier to handle in PLS (Andreev et al., 2009; Gefen et al., 2011; Hair et al., 2011).

construct (Henseler et al., 2009, p. 301). Various statistical tests can be performed to determine whether an indicator should be included in the formative construct or not (e.g., Diamantopoulos et al., 2008; Diamantopoulos and Winklhofer, 2001; Götz et al., 2010; Henseler et al., 2009; Petter et al., 2007; Urbach and Ahlemann, 2010), including: assessing the degree of multicollinearity and assessing indicators' weights as well as loadings.

High multicollinearity could mean that a formative indicator's information is redundant (Henseler et al., 2009). That is, the existence of multicollinearity may suggest that the "census" of the formative indicators was not performed successfully since formative indicators should represent distinct characteristics of the content domain, and high multicollinearity might mean that formative indicators are redundant or significantly overlapping (Andreev et al., 2009, p. 6).

In order to check for multicollinearity, variance inflation factors (VIF) were calculated (e.g., Götz et al., 2010; Henseler et al., 2009; Petter et al., 2007; Urbach and Ahlemann, 2010)<sup>14</sup>. The VIF is calculated as the inverse of the tolerance value (Black et al., 2006). Tolerance is  $1 - r^2$ , where  $r^2$  is the multiple  $r$  of a given indicator, regressed on all other indicators of the same construct. A rule of thumb from econometrics states that VIFs greater than 10 reveal a critical level of multicollinearity (e.g., Diamantopoulos and Siguaw, 2006; Gefen et al., 2000; Gefen et al., 2011; Götz et al., 2010; Henseler et al., 2009; Petter et al., 2007)<sup>15</sup>. Many IS researchers too consider VIFs up to 10 acceptable (e.g., Diamantopoulos and

Siguaw, 2006; Gefen et al., 2011; Gefen et al., 2000; Götz et al., 2010; Gujarati 2003; Henseler et al., 2009; Petter et al., 2007, Rabaa'i and Gable, 2012).

To obtain VIF and tolerance scores, several ordinary least squares (OLS) regressions were performed, with the first-order formative indicators as the independent variables and an index score for each dimension (i.e., reliability, responsiveness, rapport and tangibles) as the dependent variable. <Table 3> displays the results. The largest VIF in <Table 3> is 5.6, suggesting multicollinearity is not affecting the data in this sample.

Formative indicator weights explain the amount of variance in the formative construct explained by the indicator; a high indicator weight suggesting the indicator makes a substantive, relative contribution to the formative construct (Diamantopoulos, 2006). In PLS, the significance of formative indicator weights can be determined by means of bootstrapping (e.g., Chin, 1998; Davison and Hinkley, 2003; Götz et al., 2010; Henseler et al., 2009; Tenenhaus et al., 2005; Urbach and Ahlemann, 2010). A bootstrap analysis was performed with 200 subsamples and path coefficients were re-estimated using each of these samples. These results are also presented in <Table 4>.

From <Table 4>, four indicators (bolded) had non-significant weights. Though this might lead one to conclude that these indicators have no unique relationship with the formative construct, thereby permitting their exclusion from the model, MacKenzie et al. (2005, p. 712) state that "*dropping a measure from a formative-indicator model may omit a unique part of the conceptual domain and change the meaning of the variable, because the construct is a composite of all the indicators.*" Also, removing a non-significant formative indicator will remove the beta weight associated with it, regardless of how large or small it might be (Petter et al., 2007, p. 627).

14) The VIF indicates how much of an indicator's variance is explained by the other indicators of the same construct (Urbach and Ahlemann, 2010)

15) A VIF of 10 implies a Tolerance of 0.10 meaning that 90% of the variance in the item is explained by the other items. This of course also means 10% of the variance in the item isn't explained by the other items.

Given that each item explains only 'part' of the formative construct, formative indicator weights are typically smaller than their loadings or the loadings of reflective indicators (Götz et al., 2010). While, reflective indicators with small loadings are frequently omitted from reflective models (e.g., Götz et al., 2010, p. 698), Diamantopoulos and Winklhofer (2001) suggest "*indicator elimination - by whatever means - should not be divorced from conceptual consideration when a formative measurement model is involved*" (p. 273); it is important to ensure that the construct is still measuring the entire domain and content validity is preserved (Diamantopoulos and

Winklhofer, 2001; Jarvis et al., 2003; Petter et al., 2007).

Further, the IS service quality construct consists of 18 indicators. This large number of formative indicators has "*important implications for the statistical significance and the magnitude of each indicator's weight*" (Cenfetelli and Bassellier, 2009, p. 694). A larger number of formative indicators increases the likelihood that some of the indicator weights will be low in magnitude and some statistically non-significant (Cenfetelli and Bassellier, 2009, p. 694). Cenfetelli and Bassellier (2009, p. 695) argue that: "*formative indicators essentially "compete" with one*

<Table 4> VIFs, Weights, Loadings and Significance of the First-Order Formative ISservice Quality Indicators

IS Service Quality Indicators	VIF	Weight	Significance	Loading
<b>Reliability Indicators</b>				
Providing services as promised	2.31	0.265	$p < 0.05$	0.724
Dependability in handling user's service problems	2.65	0.342	$p < 0.001$	0.879
Performing service right the first time	1.98	0.389	$p < 0.05$	0.693
Providing services at the promised time	1.78	0.432	$p < 0.001$	0.884
Maintaining reliable technology and system	4.64	0.287	$p < 0.05$	0.622
Prompt service to users	5.38	0.103	ns	0.637
<b>Responsiveness Indicators</b>				
Willingness to help users	5.91	0.170	$p < 0.01$	0.583
Readiness to respond to user's requests	2.39	0.163	$p < 0.05$	0.616
<b>Rapport Indicators</b>				
Making users feel safer in computer transactions	2.84	0.218	$p < 0.05$	0.513
IS employees who are consistently courteous	5.91	0.106	ns	0.652
IS employees who have the knowledge to answer users' questions	5.42	0.149	$p < 0.05$	0.714
Giving users individual attention	3.84	0.458	$p < 0.001$	0.786
IS employees who deal with users in a caring fashion	5.21	0.092	ns	0.608
Having the user's best interest at heart	6.53	0.102	ns	0.548
IS employees who understand the needs of users	3.22	0.284	$p < 0.05$	0.719
<b>Tangibles indicators</b>				
Visually appealing facilities	2.46	0.335	$p < 0.001$	0.843
IS employees who appear professional	2.84	0.178	$p < 0.01$	0.648
Useful support materials (such as documentation, training, videos, etc...)	1.52	0.421	$p < 0.001$	0.925



another to be explanatory of their targeted construct. In this competition to explain variance, only a limited number of indicators will likely be significant while the others will be nonsignificant.”

From <Table 3>, we further observe that all indicators have high loadings<sup>16)</sup> (i.e., zero-order bivariate correlation) on the overall IS service quality construct. Cenfetelli and Bassellier (2009, p. 697) suggest that “it is also possible to evaluate the ‘absolute’ importance of an indicator to its construct. This is provided by the loading of the indicator and so its bivariate correlation with the formatively measured construct.” High loadings for non-significant formative indicators (bolded) suggests that although the unique contribution of each of these non-significant indicators to IS service quality construct is small in comparison to significant ones, there are still strong zero-order bivariate correlations between these non-significant indicators and the IS service quality construct.

In summary, since multicollinearity is unlikely to be a cause of low indicators’ weights and all indicators have acceptable loadings, we argue that the low weights are an artefact of the large number of indicators used to assess IS service quality, and in particular the ‘reliability’ (6 items) and ‘rapport’ (7 items) constructs with which all four non-significant items are associated.

#### 4.2. Assessment of Second-Order ‘Formative’ IS Service Quality

Following the approach of Rai et al.<sup>17)</sup> (2006),

16) It should be noted that PLS reports zero-order bivariate correlations as loadings in tandem with the weights (e.g., Cenfetelli and Bassellier, 2009).

17) This technique was also used by Bagozzi and Edwards (1999), Law and Wong (1999) and Edwards (2001b).

linear composites were computed from the indicators used to measure each of the themes in the first-order<sup>18)</sup> (i.e., reliability, responsiveness, rapport and tangibles), and used as formative indicators for the second-order IS service quality construct. To estimate second-order formative indicator weights (i.e., path coefficients), a bootstrap analysis was performed with 200 subsamples and path coefficients were re-estimated using each of these samples; results are presented in <Table 5>.

All VIFs are less than 2.65, suggesting multicollinearity is not affecting the IS service quality data in this sample. All second-order formative indicators have strong and significant path weights. Additionally, all second-order formative indicators show high loadings (i.e., zero-order bivariate correlation) on the IS service quality construct.

#### 4.3. Assessment at the Construct Level

Construct validity refers to the wider, out of the construct, validation of its measures (Straub et al., 2004). For instance, construct validity is concerned with whether or not indicators of the construct indeed measure what they intend to from the perspective of relationships between constructs, and between constructs and their relative indicators (Andreev et al., 2009, p. 6). In this study, construct validity of the IS service quality formative construct is assessed in terms of nomological validity (e.g., Andreev, et al., 2009; Henseler et al., 2009; Urbach and Ahlemann, 2010).

A nomological network includes a (i) theoretical framework of research objects, (ii) an empirical

18) Put differently, the multivariate means of the manifest variables of the first-order dimensions are used as formative indicators for the second-order IS service quality construct (Rai et al., 2006).

framework of how these objects will be measured, and (iii) specification of the relationships between these two frameworks (Campbell and Fiske, 1959). Assessing nomological validity involves evaluating the extent to which the formative construct behaves as expected within a net of hypotheses (Diamantopoulos and Winklhofer, 2001; Henseler et al., 2009; Urbach and Ahlemann, 2010). Accordingly, those relationships between the formative construct and other of the structural model constructs, which have been sufficiently referred to in prior literature, should be strong and significant (Andreev et al., 2009; Diamantopoulos and Winklhofer, 2001; Henseler et al., 2009; Straub et al., 2004; Urbach and Ahlemann, 2010). That is, testing the nomological validity of a formative construct involves (Andreev et al., 2009, p. 8):

- First, linking the focal construct with its hypothesized antecedents and consequence constructs, and
- Second, evidencing nomological validity where the hypothesized linkages (structural paths) between the constructs are found to be significantly greater than zero and their signs are in the expected direction.

The nomological validity of the IS service quality construct was assessed by linking IS service quality with the satisfaction construct in the nomological net. <Figure 1> illustrates results of the PLS analysis, showing that the relationship between the IS service

quality construct and Satisfaction construct is strong ( $\beta = 0.698, p < 0.001$ ) and significant ( $t\text{-value} = 16.23$ ), thereby evidencing the nomological validity of the IS service quality construct. Further,  $R^2$  for the satisfaction construct of 54.7% signifies that a significant and substantial part of the variance in “satisfaction” is explained by the IS service quality construct.

## V. Discussion

We return to our research questions: 1) is the overall set of IS-ServQual indicators effective in predicting customer satisfaction with the ISF?; 2) do each of the individual items make a unique contribution to the predictive power of IS-ServQual?; and 3) what (if any) is the role of the sub-dimensions (themes) of IS-ServQual? As regards the first question, we note that in nomological model testing, IS service quality has a large beta, explaining over half the variance in the satisfaction construct; thus evidencing the predictive power of the ServQual indicators as a set.

As regards the second question, the majority of the IS-ServQual indicators demonstrate significant and unique contribution to the overall IS-ServQual construct with a low level of multicollinearity between the indicators. Those items that are non-significant, nonetheless have large loadings and offer some (small

<Table 5> VIFs, Weights, Loadings and Significance of the Second-Order Formative IS Service Quality Indicators

IS Service Quality Indicators	VIF	Weight	Significance	Loading
Reliability	1.78	0.176	$p < 0.05$	0.664
Responsiveness	2.44	0.285	$p < 0.01$	0.769
Rapport	2.65	0.218	$p < 0.05$	0.771
Tangibles	2.23	0.399	$p < 0.001$	0.855

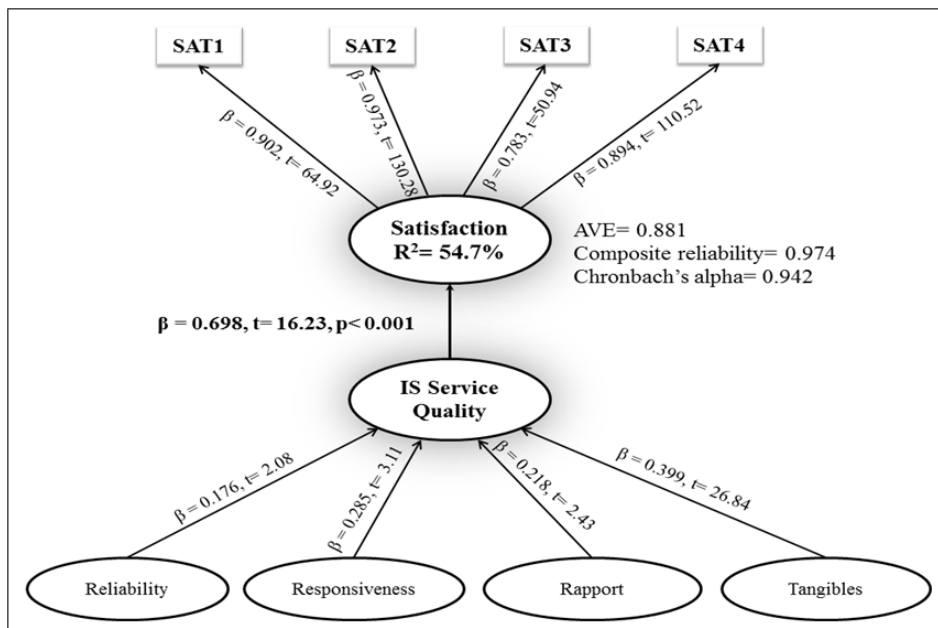
beta) contribution to the index; that contribution potentially larger in another study context. Our interpretation thus is that each item makes a 'useful' contribution and should be retained and analysed for practical implications

We note that with a few exceptions, each of the items has made a significant individual contribution to the higher-order constructs. This suggests that the items should not be evaluated reflectively as interchangeable measures. The overall means and standard deviations of the items vary greatly. This suggests that respondents are fully able to distinguish between the items in a meaningful way and that consideration of individual item scores can be effective in diagnosing service breakdowns. If we consider the "reliability" items, we can see that the mean scores for "maintaining reliable technology and system" (at 3.25) and "dependability in handling user's service problems" (at 3.51) are much less favourable than

the scores for the other items. This points to underlying problems with the IT service delivery infrastructures in the target institutions, rather than with IT staff.

Thus, the contribution to knowledge offered by the IS-ServQual instrument is best understood at the indicator/item level. Collectively, the set of indicators has considerable predictive power for user/customer satisfaction, a key dependent variable in information systems (and marketing) research. Four items had a non-significant effect in our study. However, by retaining all items, IS-ServQual findings remain more commensurable across such studies. Whether the full 18-item set is significant in any given study context is a theoretical finding of interest, rather than a defect in the instrumentation.

This discussion resolves the paradox of IS-ServQual research; that an apparently flawed measurement instrument should be so enduringly popular and



<Figure 1> Nomological Validity of the IS Service Quality Construct

influential. We suggest that the ongoing heated discourse around the stability of the ServQual and IS-ServQual dimensions has been a distraction from the real contribution of IS-ServQual research. The dimensions have been repeatedly mis-specified as reflective, when they are more appropriately viewed as themes formed from a set of organizational competencies that influence customer satisfaction.

In attention to our third question, we also validated the themes for their independent contribution to the second order formative service quality index. As evidenced in <Table 4>, each of the themes makes a unique and significant contribution to the overall measurement of IS service quality. All loaded highly on the second order construct, and they did not appear to be overly overlapping or redundant. Individually, although the themes are viewed as formative, having non-synonymous indicators and being able to vary independently, it appears that the four “themes” of reliability, responsiveness, rapport and tangibles each has some commonality which makes a distinct contribution to our overall understanding of this complex phenomenon.

From a practice perspective, we note that the 18 IS-ServQual items are highly granular and offer the possibility of providing actionable insights and interventions. The set of items has strong predictive power on satisfaction with individual item scores offering insight. Mean scores offer some indication of the relative performance of the service on each aspect to which the individual items pertain. Though the individual items are largely intuitive, each offers potential for unique interpretation in any given study/application context. In example, a low score on “IS employees who have the knowledge to answer users’ questions” should encourage reflection on who is dealing with user questions?; what is the nature of these questions?; what experience, education and

training to the IS staff have in relation to these problem areas?; are we using the right people in support?; do they need more training?; are the questions recurring?; is it possible to solve the problem so that the questions go away?; and so on.

Standard deviations have value as indicators of the level of consensus amongst respondents; strong consensus perhaps suggesting a common experience; weak consensus possibly suggesting differing experiences of different cohorts which should be explored; perhaps suggesting value in seeking to emulate the better experiences of some with those who’ve had a poorer experience. Weights indicate the relative ‘importance’ of each item to the respondents’ overall experience of service quality. Here again, as with standard deviations, there may be value in exploring the existence of sub-groups of respondents or cohorts with differential views.

### 5.1. Limitations and Further Research

While this study provides strong empirical and theoretical evidence for IS-ServQual as a multi-dimensional formative construct, it does have several limitations.

This study suffers from some of the potential issues associated with the use of formative measures, i.e., that formative construct loadings and weights tend to vary somewhat across studies/samples (Kim et al., 2010). Though four items<sup>19)</sup> are non-significant, suggesting that in this sample they make a relatively small contribution to the construct variance, or in other words, that in this sample they are largely

---

19) Three of the four non-significant items are all related to the posited “rapport” dimensions, perhaps suggesting problems with the merging of “empathy” and “assurance.” However, since we are arguing that the reflective specification of the dimensions is flawed, this finding is not of major importance for this study.

redundant; even given this, excluding them would change the definition of their dimension and the construct. And while this cross-study variation makes cross-study comparisons difficult and results that may be more sample specific, in practice this context-specificity may be valued. One might suggest that we, in this way, err in favour of internal validity over external validity, while taking greater account of context.

This study has employed an admittedly opportunistic sample in a unique context, unlike any previously reported. Further, the study has somewhat arbitrarily adopted the most recent IS-ServQual instrument published in a leading IS journal from amongst a range of alternatives employing various item sets and dimension configurations. Though consistent with the parent study from which it derives, the instrument gauges the service quality of the ISF, rather than any specific information system, as is the focus of other instrument variants. All of which suggests possible strong concerns with the generalizability of study findings. We argue, however, that the study context and instrument variant are largely irrelevant, and leave to other researchers to replicate our findings. Quantitative evidence reported herein, though constituting a substantive portion of the paper, is really supplementary to the conceptual argument presented; a conceptual argument that extends to most previously published IS-ServQual variants and their study contexts.

Though the study, data collection adopted the Kettinger and Lee (2005)'s 18 item set which has been adequate to serve our purposes, we counsel caution and propose that future studies error in favour of inclusiveness, as items dropped by Kettinger and Lee (2005) may have unique relevance in other study contexts. Some early attention to face and content validity is warranted irrespective of context. Future

researchers may even wish to revisit the original 35 items proposed by Pausaraman et al. (1985).

We have employed Kettinger and Lee (2005)'s instrument to demonstrate the statistical validity of the four dimensions when conceived as 2nd-order formative. Ultimately, given the pragmatic and predictive (rather than explanatory) value of the items, various configurations of items into dimensions may have value in practice and research; attention to such configurations demanding closer attention in future to relevance (the analytic theory qualities of utility and intuitiveness).

We note that we did not specify criterion variables for the dimensions in an effort to capture "overall" rapport. This is because we consciously chose not to reify the dimensions as distinct, measurable psychometric states which exert causal force on their indicators. This is arguably the mistake made by previous ServQual researchers. If this were possible and useful, it would almost certainly have emerged from the many reflectively specified studies already conducted. We make no claims to the independent existence or ontology of the themes, preferring to see them as indices of a set of indicators that share a common perspective on an aspect of the phenomenon.

Additionally, we note that although the IS-ServQual "dimensions" in our study appeared to make a unique and significant contribution to the overall measurement of IS-service quality and to be neither overlapping nor redundant, we counsel caution here also. Different sets of indicators might represent different conceptual dimensions of interest in different contexts. We evaluated the "dimensions" proposed by Kettinger and Lee (2005) to provide an analysis that is commensurate with their IS-ServQual study, and (as much as possible) with the original ServQual instrument. However, we acknowledge that different

composites could also be constructed.

In the study model, IS Service Quality explained 55% of satisfaction, leaving 45% unexplained. Considering that user satisfaction with ISF is a complex phenomenon with many antecedents, and IS-ServQual has high explanatory power. There is value in further research that addresses other antecedents, and ultimately, overall user satisfaction with the IS function.

## VI. Conclusion

In conclusion, we believe that our study resolves an enduring paradox, in a way that happily, allows all sides to claim some degree of success. Jiang et al. (2002) suggested that the managerial diagnostic capabilities of IS-ServQual are compelling but something that can only be accessed at the cost of empirical concerns about the validity of the measure. Our study

suggests that this is untrue, and arises from the ongoing mis-specification of the sub-dimensions of IS-ServQual as reflective. When re-specified as a multi-dimensional formative measure, IS-ServQual emerges as a high-quality measure that captures a number of important themes in the complex area of the measurement of customer satisfaction. Further, it provides a parsimonious set of indicators that have stood the test of time, and which in combination have high predictive power on user satisfaction with the ISF.

This may be symptomatic of a wider issue with our preoccupation with the identification of generalizable “constructs” and dimensions, even when the exercise has proved to be intractable, as has been the case with service quality. For too long perhaps, we have been leaning back and squinting trying to see a ‘big picture’ (the “forest”) in the pattern of evidence, when all along, it’s the individual “trees” or items that really matter.

## <References>

- [1] Albers, S. (2010). PLS and Success Factor Studies in Marketing. In V. Esposito Vinzi, H. Wang, J. Henseler and W. Chin (Eds.), *Handbook of Partial Least Squares: Concepts, Methods and Applications in Marketing and Related Fields*. Heidelberg: Springer.
- [2] Ali, A., Tate, M., Rabaa'i, A., and Zhang, M. (2012). *A Critical Evaluation and Comparison of two Formative Measures of System Quality Using Criterion Variables*. Paper presented at the 23rd Australasian Conference on Information Systems, 3-5 December 2012, Geelong, Australia.
- [3] Anderson, E. W., and Sullivan, M. W. (1993). The Antecedents and Consequences of Customer Satisfaction for Firms. *Marketing Science*, 12(2), 125-143.
- [4] Andreev, P., Heart, T., Maoz, H., and Pliskin, N. (2009). *Validating Formative Partial Least Squares (PLS) Models: Methodological Review and Empirical Illustration*. Paper presented at the 13th International Conference on Information Systems (ICIS), Phoenix, Arizona.
- [5] Au, N., Ngai, E. W. T., and Cheng, T. C. E. (2008). Extending the Understanding of End User Information Systems Satisfaction Formation: An Equitable Needs Fulfillment Model Approach. *MIS Quarterly*, 32(1), 43-66.
- [6] Au, N., Ngaib, E., and Chengb, E. (2002). A Critical Review of End-User Information System Satisfaction Research and a New Research Framework. *Omega*, 30, 451-478.
- [7] Babakus, E., and Boller, G. W. (1992). An Empirical Assessment of the SERVQUAL Scale. *Journal of Business Research*, 24(3), 253.
- [8] Bagozzi, R. (2011). Measurement and Meaning in

- Information Systems and Organizational Research: Methodological and Philosophical Foundations. *MIS Quarterly*, 35(2), 261-292.
- [9] Bagozzi, R. P., and Edwards, J. R. (1999). A General Approach for Representing Constructs in Organizational Research. *Organizational Research Methods*, 1(1), 45-87.
- [10] Bailey, E., and Pearson, W. (1983). Development of a Tool for Measuring and Analyzing Computer User Satisfaction. *Management Science*, 29(5), 530-545.
- [11] Baroudi, J., and Orlikowski, J. (1988). A Short-form Measure of User Information Satisfaction: A psychometric Evaluation of Notes on Use. *Journal of Management Information Systems*, 4(4), 44-59.
- [12] Bhattacharjee, A. (2001). Understanding Information Systems Continuance: An Expectation-Confirmation Model. *MIS Quarterly*, 25(3), 351-370.
- [13] Bhattacharjee, A., and Premkumar, G. (2004). Understanding Changes in Belief and Attitude toward Information Technology Usage: A Theoretical Model and Longitudinal Test. *MIS Quarterly*, 28(2), 229-254.
- [14] Black, J. H., W. C., Babin, B. J., Anderson, R., and Tatham, R. (2006). *Multivariate Data Analysis* (6th ed.). Upper Saddle: Prentice Hall.
- [15] Bollen, K. A. (1984). Multiple Indicators: Internal Consistency or No Necessary Relationship? *Quality and Quantity*, 18(4), 377.
- [16] Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York, NY: Wiley.
- [17] Borsboom, D. (2005). *Measuring the Mind: Conceptual Issues in Contemporary Psychometrics*. Cambridge, England: Cambridge University Press.
- [18] Borsboom, D., Mellenbergh, G. J., and van Heerden, J. (2003). The Theoretical Status of Latent Variables. *Psychological Review*, 110(2), 203-219.
- [19] Brady, M. K., Knight, G. A., Cronin, J. J. J., Hult, G. T. M., and Keillor, B. D. (2005). Removing the Contextual Lens: A Multinational, Multi-Setting Comparison of Service Evaluation Models. *Journal of Retailing*, 81(3), 215-230.
- [20] Briggs, R. O., Reinig, B. A., and de Vreede, G.-J. (2008). The Yield Shift Theory of Satisfaction and Its Application to the IS/IT Domain. *Journal of the Association for Information Systems*, 9(5), 267-293.
- [21] Brislin, R. W. (1970). Back-Translation for Cross-Cultural Research. *Journal of cross-cultural psychology*, 1, 185-216.
- [22] Brislin, R. W. (1986). The Wording and Translation of Research Instruments. In W. Lonner and W. Berry (Eds.), *Field Method in Cross-culture Research* (pp. 137-164). Beverly Hills, CA: Sage.
- [23] Campbell, D. T., and Fiske, D. W. (1959). Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin*, 56(2), 81-105.
- [24] Carman, J. (1990). Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions. *Journal of Retailing*, 66(1), 33-55.
- [25] Carr, L. (2002). A Psychometric Evaluation of the Expectations, Perceptions, and Difference-Scores Generated by the IS-Adapted SERVQUAL Instrument. *Decision Sciences*, 33(2), 281-296
- [26] Cenfetelli, R. T., and Bassellier, G. (2009). Interpretation of Formative Measurement in Informations Systems Research. *MIS Quarterly*, 33(4), 689-707.
- [27] Cenfetelli, R. T., Benbasat, I., and Al-Natour, S. (2008). Addressing the What and How of Online Services: Positioning Supporting-Services Functionality and Service Quality for Business-to-Consumer Success. *Information Systems Research*, 19(2), 161-181.
- [28] Chin, W. (1998). The Partial Least Squares Approach to Structural Equation Modeling. In G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 295-358). Mahwah, NJ: Lawrence Erlbaum Associates.
- [29] Davison, A. C., and Hinkley, D. V. (2003). *Bootstrap Methods and their Application* (2nd ed.). New York: Cambridge University Press.
- [30] Delone, H. (1988). Determinants of Success for Computer Usage in Small Business. *MIS Quarterly*, 12(1), 51-61.
- [31] Diamantopoulos, A. (1999). Export Performance Measurement: Reflective Versus Formative Indicators. *International Marketing Review*, 16(6), 444-457.
- [32] Diamantopoulos, A. (2006). The Error Term in Formative Measurement Models: Interpretation and

- Modeling Implications. *Journal of Modelling in Management*, 1(1), 7-17.
- [33] Diamantopoulos, A., Riefler, P., and Roth, K. P. (2008). Advancing Formative Measurement Models. *Journal of Business Research*, 61(12), 1203-1218.
- [34] Diamantopoulos, A., and Sigauw, J. A. (2006). Formative Versus Reflective Indicators in Organizational Measure Development: A Comparison and Empirical Illustration. *British Journal of Management* 17(4), 263-282.
- [35] Diamantopoulos, A., and Winklhofer, H. M. (2001). Index Construction with Formative Indicators: An Alternative to Scale Development. *Journal of Marketing Research*, 38(2), 269-277.
- [36] Diamantopoulos, A., and Winklhofer, H. (2001). Index Construction with Formative Indicators: an Alternative to Scale Development. *Journal of Marketing Research*, 38(2), 269-277.
- [37] Doll, W. J., and Torkzadeh, G. (1988). The Measure of End-User Computing Satisfaction. *MIS Quarterly*, 12(2), 259-274.
- [38] Edwards, J. R. (2001a). Multidimensional Constructs in Organisational Behavior Research: An Integrative Analytical Framework. *Organisational Research Methods*, 4(2), 144-192.
- [39] Edwards, J. R. (2001b). Multidimensional Constructs in Organizational Behavior Research: An Integrative Analytical Framework. *Organizational Research Methods*, 4(2), 144-192.
- [40] Edwards, J. R., and Bagozzi, R. P. (2000). On the Nature and Direction of Relationships between Constructs and Measures. *Psychological Methods*, 5(2), 155-174.
- [41] Etezadi-Amoli, J., and Farhoomand, A. F. (1996). A Structural Model of End User Computing Satisfaction and User Performance. *Information and Management*, 30(2), pp. 65-73.
- [42] Fisk, R., Brown, S., and Bitner, M. (1993). Tracking the Evolution of the Services Marketing Literature. *Journal of Retailing*, 69(1), 61-103.
- [43] Gable, G., and Sedera, D. (2009). *Formative and Reflective Measurement and Validation Mismatch in Survey Research : an Archival Analysis of Information Systems Success Constructs 1985-2007*. Paper Presented at the International Conference on Information Systems (ICIS), 15-18 December, Phoenix, Arizona.
- [44] Gatian, A. W. (1994). Is User Satisfaction a Valid Measure of System Effectiveness. *Information and Management* 26(3), 119-131.
- [45] Gefen, D., Rigdon, E., and Straub, D. (2011). EDITOR'S COMMENTS: An Update and Extension to SEM Guidelines for Administrative and Social Science Research. *MIS Quarterly*, 35(2), iii-xiv.
- [46] Goodhue, D. (1995). Understanding User Evaluations of Information Systems. *Management Science*, 41(12), 1827-1844.
- [47] Gorla, N., Somers, T., and Wong, B. (2010). Organizational Impact of System Quality, Information Quality and Service Quality. *Journal of Strategic Information Systems*, 2010(19), 207-228.
- [48] Götz, O., Liehr-Gobbers, K., and Krafft, M. (2010). Evaluation of Structural Equation Models Using the Partial Least Squares (PLS) Approach. In V. Esposito Vinzi, H. Wang, J. Henseler and W. W. Chin (Eds.), *Handbook of Partial Least Squares: Concepts, Methods and Applications in Marketing and Related Fields*. Heidelberg: Springer.
- [49] Hair, J. F., Ringle, C. M., and Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice*, 19(2), 139 - 151.
- [50] Hansen, C. (1987). Cross-Cultural Research on Vocational Interests. *Measurement and Evaluation in Counseling and Development*, 19(4), 163-176.
- [51] Henseler, J., Ringle, C., and Sinkovics, R. (2009). The use of Partial Least Squares Path Modeling in International Marketing. *Advances in International Marketing*, 20, 277-320.
- [52] Hughey, D., Chawla, S., and Khan, Z. (2003). Measuring the Quality of University Computer Labs using SERVQUAL: a Longitudinal Study. *The Quality Management Journal*, 10(3), 33-44.
- [53] Igarria, M., and Nachman, A. (1990). Correlates of User Satisfaction With End User Computing: An Exploratory Study. *Information and Management* 19(2), 73-82.



- [54] Igbaria, M., and Tan, M. (1997). The Consequences of Information Technology Acceptance on Subsequent Individual Performance. *Information and Management*, 32(3), 113-121.
- [55] Iivari, J. (1987). *User information satisfaction (UIS) reconsidered: An information system as the antecedent of UIS*. Paper presented at the 8th International Conference on Information Systems (ICIS).
- [56] Iivari, J. (1997). User Information Satisfaction: A Critical Review. *Encyclopedia of Library and Information Science*, 60, 341-364.
- [57] Ives, B., Olson, H., and Baroudi, J. (1983). The Measurement of User Information Satisfaction. *Communications of the ACM*, 26(10), 785-793.
- [58] Jarvis, C. B., MacKenzie, S. B., and Podsakoff, P. M. (2003). A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research. *Journal of Consumer Research*, 30(2), 199-218.
- [59] Jayawardhena, C. (2004). Measurement of Service Quality in Internet Banking: The Development of an Instrument. *Journal of Marketing Management*, 20(1/2), 185-207
- [60] Jia, R., and Reich, B. H. (2013). IT service climate, antecedents and IT service quality outcomes: Some initial evidence. *The Journal of Strategic Information Systems*, 22(1), 51-69.
- [61] Jia, R., Reich, B. H., and Pearson, J. M. (2008). IT Service Climate: An Extension to IT Service Quality Research. *Journal of the Association for Information Systems*, 9(5), 294-320.
- [62] Jiang, J., Klein, G., and Carr, C. (2002). Measuring Information System Service Quality: SERVQUAL from the Other Side. *MIS Quarterly*, 26(2), 145-166.
- [63] Kettinger, W. J., and Lee, C. C. (1995). Exploring a "Gap" Model of Information Services Quality. *Information Resources Management Journal*, 8(3), 5-16.
- [64] Kettinger, W. J., and Lee, C. C. (1997). Pragmatic Perspectives on the Measurement of Information Systems Service Quality. *MIS Quarterly*, 21(2), 195-208.
- [65] Kettinger, W. J., and Lee, C. C. (2005). Zones of Tolerance: Alternative Scales for Measuring Information Systems Service Quality. *MIS Quarterly*, 29(4), 607-623.
- [66] Khalifa, M., and Liu, V. (2004). The State of Research on Information System Satisfaction. *Journal of Information Technology Theory and Application*, 5(4).
- [67] Klein, G., Jiang, J., and Cheney, P. (2009). Resolving Difference Score Issues in Information Systems Research. *MIS Quarterly*, 33(4), 811-826.
- [68] Klenke, K. (1992). Construct Measurement in Management Information Systems: A review and Critique of User Satisfaction and User Involvement Instruments. *Information and Management*, 30(4), 325-348.
- [69] Law, K. S., and Wong, C. (1999). Multidimensional Constructs In Structural Equation Analysis: An Illustration Using the Job Perception and Job Satisfaction Constructs. *Journal of Management Information Systems*, 25(2), 143-160.
- [70] Law, K. S., Wong, C., and Mobley, W. H. (1998). Toward a Taxonomy of Multidimensional Constructs. *Academy of Management Review*, 23(4), 741-755.
- [71] Li, Y., Tan, K., and Xie, M. (2003). Factor Analysis of Service Quality Dimension Shifts in the Information Age. *Computing Surveys*, 10(1), 19-34
- [72] MacCallum, R. C., and Browne, M. W. (1993). The use of Causal Indicators in Covariance Structure Models: Some Practical Issues. *Psychological Bulletin*, 114, 533-541.
- [73] MacKenzie, S., Podsakoff, P., and Podsakoff, N. (2011). Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques. *MIS Quarterly*, 35(2), 293-334.
- [74] MacKenzie, S. B., Podsakoff, P. M., and Jarvis, C. B. (2005). The Problem of Measurement Model Misspecification in Behavioral and Organizational Research and Some Recommended Solutions. *Journal of Applied Psychology*, 90(4), 710-730.
- [75] Mahmood, M. A., and Becker, J. D. (1985). Effect of Organizational Maturity on End-Users' Satisfaction with Information Systems. *Journal of Management Information Systems*, 2(3), 37-64.

- [76] Mels, G., Boshoff, C., and Nel, D. (1997). The Dimensions of Service Quality: The original European perspective revisited. *The Service Industries Journal*, 17(1), 173-189.
- [77] Netemeyer, R. G., Bearden, W. O., and Sharma, S. (2003). *Scaling Procedures: Issues and Applications*. Thousand Oaks: Sage Publications.
- [78] O'Cass and Carlson (2012). An Empirical Assessment of Consumers' Evaluations of Web Site Service Quality: Conceptualizing and Testing a Formative Model. *Journal of Service Marketing*, 26(6), 419-434
- [79] Oliver, R. (1981). Measurement and Evaluation of Satisfaction Processes in Retail Settings. *Journal of Retailing*, 57(3), 25-48.
- [80] Oliver, R. (1993). Cognitive, Affective, and Attribute Bases of the Satisfaction Response. *Journal of Consumer Research*, 20(3), 418-430.
- [81] Oliver, R. (1997). *Satisfaction : A Behavioral Perspective on the Consumer*. New York: McGraw Hill.
- [82] Oliver, R., and Swan, J. (1989). Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions: A Field Survey Approach. *Journal of Marketing*, 53, 21-35.
- [83] Parasuraman, A., Berry, L., and Zeithaml, V. (1991). Refinement and Reassessment of Service Quality. *Journal of Retailing*, 64(1), 12-40.
- [84] Parasuraman, A., Berry, L., and Zeithaml, V. (1991). Understanding Customer Expectations of Service. *Sloan Management Review*, 32(3), 39-48.
- [85] Parasuraman, A., Zeithaml, V., and Berry, L. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, 49(4), 41-50.
- [86] Parasuraman, A., Zeithaml, V., and Berry, L. (1988). SERVQUAL: A Multi-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing*, 64(Spring), 12-40.
- [87] Petter, S., Straub, D., and Rai, A. (2007). Specifying Formative Constructs in Information systems Research. *MIS Quarterly*, 31(4), 623-656.
- [88] Pitt, L., Watson, R., and Kavan, B. (1995). Service Quality: A Measure of Information Systems Effectiveness. *MIS Quarterly*, 19(2), 173-187.
- [89] Po-An Hsieh, J.J., Rai, A., Petter, S. and Zhang, T. (2012). Impact of User Satisfaction with Mandated CRM Use on Employee Service Quality. *MIS Quarterly*, 36(4), 1065-1080.
- [90] Polites, G. L., Roberts, N., and Thatcher, J. (2011). Conceptualizing Models Using Multidimensional Constructs: A Review and Guidelines for Their use. *European Journal of Information Systems*, 20(1), 1-27.
- [91] Premkumar, G., and Bhattacharjee, A. (2008). Explaining Information Systems Usage: A Test of Competing Models. *Omega*, 36(1), 64-75.
- [92] Rabaa'i, A., and Gable, G. (2012). *IS Service Quality as a Multi-Dimensional Formative Construct*. Paper Presented at the 16th Pacific Asia Conference on Information Systems, 11-15 July, 2012. Ho Chi Minh city, Vietnam.
- [93] Rai, A., Patnayakuni, R., and Seth, N. (2006). Firm Performance Impacts of Digitally Enabled Supply Chain Integration Capabilities. *MIS Quarterly*, 30(2), 225-246.
- [94] Ringle, C., Wende, S., and Will, A. (2005). SmartPLS Version 2.0 M2, from <http://www.smartpls.de>
- [95] Rossiter, J. R. (2002). The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing*, 19(4), 305-335.
- [96] Samaddar, S., and Kadiyala, S. (2006). Information Systems Outsourcing: Replicating an Existing Framework in a Different Cultural Context. *Journal of Operations Management*, 24(6), 910-931.
- [97] Sedera, D., and Tan, F. (2005). *User Satisfaction: A Dimension of Enterprise System Success*. Paper presented at the 9th Pacific Asia Conference on Information Systems (PACIS), Thailand.
- [98] Spreng, R., MacKenzie, S., and Olshavsky, R. (1996). A Reexamination of the Determinants of Consumer Satisfaction. *Journal of Marketing Research*, 60(3), 15-32.
- [99] Spreng, R., and Mackoy, R. (1996). An Empirical Examination of a Model of Perceived Service Quality and Satisfaction. *Journal of Retailing*, 72(2), 201-214.
- [100] Sylvester, A., Tate, M., and Johnstone, D. (2013)

- Beyond synthesis: representing heterogeneous research literature. *Behaviour and Information Technology*, 32(12), 199-1215.
- [101] Tate, M., and Evermann, J. (2010). The End of ServQual in Online Services Research: Where to from here? *e-Service Journal*, 7(1), 60-85.
- [102] Teas, R. (1994). Expectations as a Comparison Standard in Measuring Service Quality: An Assessment of a Reassessment" *Journal of Marketing*, 58(1), 132-139.
- [103] Tenenhaus, M., Vinzi, V. E., Chatelin, Y.-M., and Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis*, 48, 159-205.
- [104] Urbach, N., and Ahlemann, F. (2010). Structural Equation Modeling in Information Systems Research Using Partial Least Squares. *Journal of Information Technology Theory and Application*, 11(2), 5-40.
- [105] Van Dyke, T. P., Kappelman, L., and Prybutok, V. (1997). Measuring Information Systems Service Quality: Concerns on the SERVQUAL Questionnaire. *MIS Quarterly*, 21(2), 195-208.
- [106] VanDyke, T., Prybutok, V., and Kappelman, L. (1999). Cautions on the use of the SERVQUAL Measure to Assess the Quality of Information Systems Services. *Decision Sciences*, 30(30), 877-891.
- [107] Vlachos, P. A., and Theotokis, A. (2009). Electronic Service Quality in Mobile Music Services: Comparing Different Second-Order Measurement Specification. In A. Poulmenakou, N. Pouloudi and K. Pramataris (Eds.), *4th Mediterranean Conference on Information Systems*. Athens, Greece.
- [108] Wang, Y., and Tang, T. (2003). Assessing Customer Perceptions of Web Sites Service Quality in Digital Marketing Environments. *Journal of End User Computing*, 15(3), 14-31
- [109] Wetzels, M., Odekerken-Schröder, G., and van Oppen, C. (2009). Using PLS Path Modeling for Assessing Hierarchical Construct Models: Guidelines and Empirical Illustration. *MIS Quarterly*, 33(1), 177-195.
- [110] Woodroof, J., and Kasper, G. (1998). A Conceptual Development of Process and Outcome User Satisfaction. *Information Resources Management Journal*, 11(2), 37-43.
- [111] Yang, Y., Stafford, T., and Gillenson, M. (2011). Satisfaction with Employee Relationship Management Systems: The Impact of Usefulness on Systems Quality Perceptions. *European Journal of Information Systems*, 20, 221-236.
- [112] Yang, Z., Jun, M., and Peterson, R. T. (2004). Measuring Customer Perceived Online Service Quality. *International Journal of Operations and Production Management*, 24(11), 1149-1174.
- [113] Zviran, M., and Erlich, Z. (2003). Measuring IS User Satisfaction: Review and Implications. *Communications of the Association for Information Systems*, 12(1), 81-103.

## ◆ About the Authors ◆

---



**Ahmad A. Rabaa'i**

Ahmad A. Rabaa'i is an Assistant Professor in the Department of Computer Science and Information Systems, at the American University of Kuwait (AUK), Kuwait. Rabaa'i has published over 20 peer-reviewed articles in conferences and journals including International Journal of Intercultural Information Management, International Journal of Information Systems and Management, and leading international and regional conferences such as PACIS and ACIS. Rabaa'i has research interests in service quality and service management, IS success and impact, cultural studies and technology, social media and research methods.

---



**Mary Tate**

Mary Tate is a Senior Lecturer in the School of Information Management, at Victoria University of Wellington, New Zealand. Tate has published over 70 peer-reviewed articles in conferences and journals including The Journal of the Association for Information Systems, Communications of the Association for Information Systems, Behaviour and Information Technology, and leading international and regional conferences such as ICIS, ECIS, PACIS and ACIS. Tate has research interests in service quality and service management, IS success and impact, individual attitudes towards technology, and research methods. Tate is on the editorial board of the Communications of the AIS.

---



**Guy Gable**

Professor Guy Gable leads the 'Research Systems' research program (>100 publications e.g., MS, JAIS, JSIS, I & M, EJIS) Information Systems School, Queensland University of Technology, Australia. He is Senior Editor Journal of Strategic Information Systems (since 2003) and has served on the editorial boards of MISQ, JAIS, EJIS and others. He is currently Adjunct Professor Renmin University of China; and a Senior Visiting Professor of Xi'an Jiaotong University's Next Generation Management and Innovation Research Center. A Charter Member of AIS, he has held various AIS roles, and has been highly active in the main regional conference 'Pacific Asia Conference on IS' (2013 recipient of the 'PACIS Plaque of Appreciation' for lifetime contribution). Australian Research Council (ARC) grants on which he is 1st Chief Investigator include two ARC Discoveries and an ARC Linkage ~2012-2016 (totalling AUD\$2.6M) in collaboration with RENMIN and XI'AN, Georgia State University, Victory University of Wellington, and University of Sydney.

---

Submitted: February 20, 2015; 1st Revision: April 25, 2015; Accepted: April 28, 2015