Association for Information Systems
AIS Electronic Library (AISeL)

ICIS 2010 Proceedings

International Conference on Information Systems (ICIS)

2010

THE ROLE OF SAAS SERVICE QUALITY FOR CONTINUED SAAS USE: EMPIRICAL INSIGHTS FROM SAAS USING FIRMS

Alexander Benlian
Ludwig-Maximilians-University Munich, benlian@ise.tu-darmstadt.de

Marios Koufaris
Baruch College, CUNY, Marios.Koufaris@baruch.cuny.edu

Thomas Hess
Ludwig-Maximilians-University Munich, thess@bwl.lmu.de

THE ROLE OF SAAS SERVICE QUALITY FOR CONTINUED SAAS USE: EMPIRICAL INSIGHTS FROM SAAS USING FIRMS

Completed Research Paper

Alexander Benlian
Institute for Information Systems and New Media
Ludwig-Maximilians-University Munich
Ludwigstr. 28, 80539 Munich, Germany
benlian@bwl.lmu.de

Marios Koufaris
SCIS Department, Zicklin School of Business, Baruch College, CUNY
55 Lexington Avenue, Box B11-220
New York, NY 10010
Marios.Koufaris@baruch.cuny.edu

Thomas Hess
Institute for Information Systems and New Media
Ludwig-Maximilians-University Munich
Ludwigstr. 28, 80539 Munich, Germany
thess@bwl.lmu.de

Abstract

Despite its success in the software industry, Software-as-a-Service (SaaS) still struggles with fulfilling customer expectations regarding service quality. To contain customer churn rates to low levels, SaaS providers have to address their service quality weak spots and find out which factors are crucial for continued SaaS usage. Drawing on previous service quality literature, we develop a Zone-of-Tolerance (ZOT)-based SaaS-QUAL scale and validate it in a model of IS continuance based on two empirical surveys of SaaS using firms. By doing this, we examine the importance of SaaS service quality factors for shaping customer satisfaction and SaaS continuance intentions. Furthermore, we provide insights into what service factors effectively meet or miss SaaS customer expectations. As a practical contribution, we develop and apply a SaaS-QUAL scale that can be used as a diagnostic tool by SaaS providers and users alike. For researchers, we enrich existing research models on IS continuance by integrating a more fine-grained conceptualization of service quality confirmation that provides stronger explanatory power than in previous models.

Keywords: Software-as-a-Service, Service quality, SERVQUAL, Satisfaction, IS continuance, IS post-adoption, Zone of tolerance

Introduction

According to a study by Gartner, Software-as-a-Service (SaaS) is predicted to be increasingly important in most enterprise application software (EAS) markets. Worldwide software revenues for SaaS delivery are forecast to grow from 2008 to 2013 by 19.4% overall, which is more than triple the total market compound annual growth rate of 5.2% (Mertz et al. 2009, Pettey 2006). After several years of diffusion, the SaaS market is already going through a maturation process. SaaS has been tried and tested through users and IS departments in small- to medium-sized and large enterprises (Benlian 2009). For software providers, software service delivery via the Internet thus becomes an increasingly relevant channel for selling software services. If these channels are to be viable, however, they must be perceived by consumers as effective and efficient.

Even though IT cost reductions, operational elasticity, faster upgrade cycles, and ease of implementation of SaaS were initially thought to be the drivers of success, service quality issues are increasingly becoming pivotal. As recent reports show, there has been an increase in instances where the provision of SaaS offerings has missed customers' service quality expectations. For example, on September 24th, 2009, Workday's on-demand ERP services for human resources, financial applications and payroll was down for 15 hours (Weier 2009). This led to Workday's customers losing the ability to process human resources transactions and affected their payment services, payroll and other important financial process. Examples such as this show that failing to fulfill customers' expectations regarding service quality, such as application availability or vendor responsiveness, may have critical consequences not only for customers, but also for the vendor. Due to shorter contract cancellation periods and lower switching costs in a SaaS-based relationship, it is much easier for customers to press the eject button and churn to another software vendor. In a Gartner study of 333 US- and UK-based organizations (Pring and Lo 2009), the top three reasons why organizations discontinue SaaS or put SaaS on hold are security/privacy issues, technical integration problems, and low-quality customer support.

Given the apparent importance of service quality aspects, the SaaS business model can't solely rely on cost and implementation advantages compared to on-premise solutions. If SaaS is to be accepted and continuously used by its costumers, SaaS vendors also need to shift the focus to all relevant aspects of service quality management – i.e. all cues and encounters that occur before, during, and after the delivery of software services. To deliver superior service quality, managers of companies with SaaS offerings must thus understand how costumers perceive and evaluate SaaS-based services. In this way they know in which areas to allocate investments to improve their service quality and to increase continued SaaS usage (i.e. SaaS renewal rates).

Although previous models on continued IS usage have examined the influence of IS service quality confirmation on satisfaction and of continued IS usage intentions (e.g. Bhattacherjee 2001; Deng et al. 2010), they used rather abstract notions of service quality, which is highly desirable for theory building purposes. However, in order to offer more diagnostic and thus prescriptive advice, a more in-depth conceptualization of SaaS service quality provides more insights into were the weak spots (and strengths) of SaaS services are that may explain dissatisfaction and possible discontinuance of SaaS usage. Integrating a richer conceptualization of SaaS service quality into existing models of IS continuance thus seems to make both a practical and conceptual contribution to the emerging service science and service management literature. In light of the growing interest in the service paradigm, Rai and Sambamurthy (2006) note that "important questions emerge on customer perceptions and the economics of digitally enabled services", and thus suggest the examination of "the economics and customer experiences with these services" (Rai and Sambamurthy 2006, p. 330).

Given this call for research and the research gap identified above, our paper aims to contribute to a more sophisticated understanding of the role of SaaS service quality factors in shaping customer satisfaction, perceived usefulness and continued SaaS usage intentions. Our research questions are:

(1) How does SaaS service quality contribute to explaining SaaS continuance intentions? Which SaaS service quality factors are more or less important in affecting key antecedents of SaaS continuance intentions?

(2) How do SaaS customers assess the actual performance of relevant SaaS service quality factors relative to their expectations? Which performance-expectation gaps should SaaS providers and customers address to improve service provision and experience?

The remainder of this article is organized as follows. The next section develops the theoretical basis for this work drawing on literature in service quality and IS continuance research. We then present the research model and related

hypotheses, describe the research methodology and the two empirical studies used to test the proposed hypotheses. After discussing the results, the paper highlights implications for both research and practice and points out limitations and promising areas for future research.

Theory Background

Previous Service Quality Research and SaaS Service Quality

There is a long tradition of research on traditional (non-Internet-based) service quality (SQ) over the past 25 years (see Parasuraman and Zeithaml 2002 for a review), beginning with the early proposals that SQ stems from a comparison of what customers feel a company should offer (i.e. their expectations) with the company's actual service performance (Lewis and Booms 1983; Parasuraman et al. 1985). Using insights from these studies as a starting point, Parasuraman et al. (1991) conducted empirical studies in several industry sectors to develop and refine SERVQUAL, a multiple-item instrument to quantify customers' global assessment of a company's SQ. This scale measures SQ along five dimensions: *reliability*, *responsiveness*, *assurance*, *empathy*, and *tangibles*.

The service quality literature and the SERVQUAL scale have been adapted to the IS context through several studies investigating the service quality of IS functions and departments from the perspective of users or IS professionals (e.g. Watson et al. 1998; Kettinger and Lee 1997; Jiang et al. 2002). Despite criticisms on conceptual and psychometric grounds (e.g. Kohlmeyer and Blanton 2000; van Dyke et al. 1999), researchers and practitioners continuously emphasized SERVQUAL's diagnostic and thus its practical relevance for management decisions (e.g. Jiang et al. 2002; Pitt et al. 1997). A primary area of criticism concerns SERVQUAL's reliance on gap scores that are derived by calculating the difference between IS users' perceived levels of service and their expectations for service. Critics both in marketing (e.g. Brown et al. 1993; Cronin Jr and Taylor 1992; Teas 1993) and in IS (e.g. van Dyke et al. 1999; van Dyke et al. 1997) have pointed to conceptual and empirical difficulties with the original SERVQUAL instrument and have suggested that alternatives to the original "gap scored" IS-adapted SERVQUAL be explored. To overcome some of the most significant points of criticisms with the original SERVQUAL instrument while retaining the practical diagnostic power of understanding (IS) service expectation levels, Kettinger and Lee (2005, 1997) tested and validated an alternative instrument adapted from marketing referred to as the "Zones of Tolerance" (ZOT) service quality measure (Zeithaml et al. 1993). This new instrument recommended using "two different comparison norms for service quality assessment: desired service (i.e. the level of service a customer believes can and should be delivered) and adequate (minimum) service (i.e. the level of service the customer considers acceptable)." Separating these two levels is a zone of tolerance that represents the range of service performance a customer would consider satisfactory. In other words, customer service expectations are characterized by a range of levels (between desired and adequate service), rather than a single point. By providing precise information about the perceived service levels across different relevant dimensions relative to adequate and desired service levels, insight is gained concerning the amount of emphasis that should be placed on improving poorly evaluated dimensions.

Through the emergence of the Internet and electronic channels, several adaptations to the IS SERVQUAL measurement scale have been proposed. Researchers found that studying online SQ requires scale development that extends beyond merely adapting offline scales. Gefen (2002), for example, extended the SERVQUAL conceptualization to the online context and found that the five service quality dimensions collapse to three: (a) *tangibles*; (b) a combined dimension of *responsiveness, reliability, and assurance*; and (c) *empathy*. On the basis of a comprehensive review of the extant literature on online SQ, Zeithaml et al. (2002) detailed five broad sets of criteria as relevant to online SQ perceptions: (a) *information availability and content*, (b) *ease of use or usability*, (c) *privacy/security*, (d) *graphic style*, and (e) *reliability/fulfillment*. Drawing on previous research in IS SERVQUAL and online SQ (E-Commerce) literature, a few IS researchers have transferred and adapted the findings to the application service provider (ASP) context. In an exploratory study, Ma et al. (2005) developed a ZOT-based ASP-QUAL scale capturing the specifics of this software business model. The study found that service quality in the realm of ASP comprises seven factors including *features, availability, reliability, assurance, empathy, conformance, and security*. Sigala (2004) developed an ASP service quality model for companies evaluating their ASP-hosted online stores. Her analysis suggested a multi-factor scale including *tangibles, reliability, responsiveness, assurance, empathy, trust, business understanding, benefit and risk share, conflict and commitment*. Since both studies focused

on the development of valid and reliable measurement scales for ASP-QUAL, they lacked the assessment of the scales' nomological validity and did not examine the relative importance of specific service quality facets.

Due to some technical and economic shortcomings of the ASP model, SaaS emerged as an alternative way to provide software services. In this new multi-tenant architecture, only a single instance of the common code and data definitions for a given application exists on the vendor's server, and no customization of this code is permitted (Menken 2008). Customer-specific configuration can only be made at the meta-data layer on top of the common code using interfaces provided by the SaaS vendor. The service can be integrated with other applications or connect with more custom functions through common web services application programming interfaces (APIs) that are defined and maintained by SaaS vendors (Xin and Levina 2008).

The new architecture has important implications for customers' service quality perceptions. While the findings in IS SERVQUAL and ASP research can be borrowed to guide research in SaaS service quality, SaaS differs in several aspects from previous models. First, contrary to the classical ASP model where software applications and IT infrastructure (i.e. servers, storage, and bandwidth) are dedicated to each single customer, they are shared across customers in the SaaS model which may have implications for system performance, availability and security/privacy aspects. Second, the SaaS model constrains clients' options for customization of the main functionality and data structures of the software. Third, it gives more control over future development to the vendor, as clients have no choice but to adopt future upgrades of software if they continue using the service because interfaces are most frequently not backward-compatible. Fourth, the architecture of SaaS allows for the separation of maintenance responsibilities between the SaaS vendor and the client. In particular, the SaaS vendor is responsible for maintaining the common code base that delivers the standard application services to all customers, while customers are responsible for maintaining their custom-developed code. Thus, this model no longer requires any client-specific investment by the vendor and helps vendors to reap significant economies of scale for they do not need to constantly keep increasing the size of their data centers. They thus can share one application cost effectively across hundreds of customers which is a vast improvement on the old ASP model (Choudhary 2007).

In sum, SaaS proponents claim that SaaS allows providers to offer customers technologically more mature and more 'modularized' service packages than the ASP model and, from a total-cost-of-ownership point-of-view, a more inexpensive access to applications via easy-to-use Internet browsers (Waters, 2005). On the other side, SaaS skeptics point out that limited customization possibilities of SaaS applications and potential traffic bottlenecks due to sharing IT infrastructure across many customers may hamper service quality dimensions. Given these specific features of SaaS and their implications for service quality, they have to be considered when examining the role of SaaS service quality for SaaS continuance intentions.

IS Continuance Research

Research in technology acceptance has primarily drawn on theories adapted from social psychology (theory of reasoned action, TRA; theory of planned behavior, TPB; and so forth) to explore the many antecedents and moderating effects leading to the initial acceptance of a particular IS (e.g. Davis 1989; Bajaj and Nidumolu 1998; Compeau et al. 1999). More recently, scholars have begun to also study the nature of IS continuance (Jasperson et al. 2005; Bhattacherjee 2001). IS continuance, IS continuance behavior, or IS continuous usage describe post-adoption behavioral patterns reflecting continued use of a particular IS. Though the term post-adoption actually refers to behaviors that follow initial acceptance, it is often used as a synonym for continuance (see Karahanna et al. 1999). In this study, we limit ourselves to the terms IS continuance or continued IS usage behavior.

IS continuance has been explored both at the organizational and individual level of analysis. Saga and Zmud (1994) associated the IS post-adoption at the organizational level with the final three phases of their six-stage IT implementation model. These phases include organizational efforts undertaken to induce organizations to commit to the use of IT (acceptance), alterations occurring within the work system such that they are no longer perceived as new (routinization), and the process of embedding IT into the organization's work system (infusion). Studied at the individual level, IS continuance behavior refers to a usage stage when IS use transcends conscious behavior and becomes part of normal routine activity (Bhattacherjee 2001). Unlike the initial adoption decision, IS continuance is not a one-time event, but the result of a series of individual decisions to continue using a particular IS, thus reflecting its longitudinal nature. The IS continuance phase ends with the users' final decision to discontinue.

Research Model and Hypotheses Development

Dismissing adoption research as largely unsuitable to explain phenomena in IS continuance, Bhattacherjee's (2001) work builds on different theoretical foundations by turning to research in consumer satisfaction. According to his post-acceptance model of IS continuance, which is based on expectation-confirmation theory (ECT), IS users' continuance decisions are similar to consumer's repurchase decisions, are influenced by the initial use, and can potentially lead to an ex-post reversal of the initial decision, that is, to the discontinuance of the IS.

While Bhattacherjee's model has been validated in several follow-up studies (e.g. Deng et al. 2010; Limayem et al. 2007; Hsu et al. 2004) and thus has proven its theoretical value, its diagnostic and thus practical value has been limited due to a rather abstract notion of service quality confirmation aspects. Bhattacherjee's and also follow-up models used only few and rather abstract items to measure the confirmation of users' expectations towards the system's performance and service levels provided. Delivering actionable guidance to practitioners for specific application settings (such as SaaS) about SQ-related causes of customers' discontinuance intentions, however, requires a more fine-grained conceptualization of service quality confirmation. To overcome these limitations and to improve the prescriptive value of the model further, we feel it needs to be extended by incorporating a more in-depth conceptualization of service quality. This way the model would not only more comprehensively embrace the rich literature in IS SERVQUAL research and examine the theoretical contribution of an enriched conceptualization of service quality confirmation. It would also be able to account for more specific facets of service quality that influence customer satisfaction and perceived usefulness and inform practitioners on the relative importance of different SaaS service quality drivers. Incorporating the ideas described in the previous sections into Bhattacherjee's (2001) model, we thus aim to enrich it by including an antecedent variable – called confirmation of SaaS service quality (see Figure 1) – that is refined and conceptually informed by the ZOT-based IS SERVQUAL and SaaS literature and should capture all relevant dimensions of SaaS service quality (i.e. the sub-dimensions that are part of SaaS service quality were identified and validated in a scale refinement process which is presented in the methodology section) in order to increase the models diagnostic value. We develop the rationale for the relationships in our research model below.

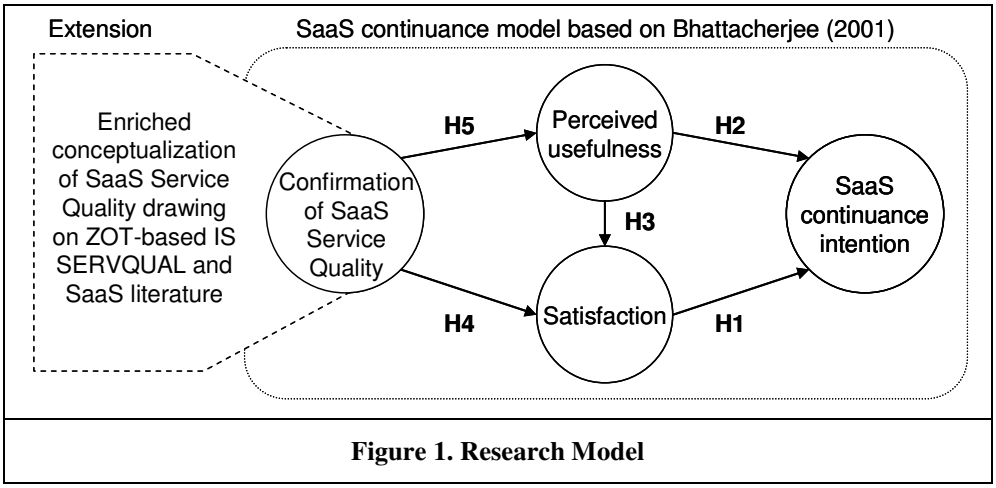

Figure 1. Research Model

In the context of our study, we argue that – consistent with Bhattacherjee's model – a similar argument can be made in the context of IS continuance that satisfaction with an IS tends to reinforce a user's intention to continue using the IS system. We expect the same argument to also apply to the continuance intention of SaaS users.

H1: *Satisfaction has a positive association with SaaS continuance intention.*

Further, by including perceived usefulness, Bhattacherjee's model reflects current thinking in the area of IS which holds that perceived usefulness is the only construct consistently influencing user intention across both adoption and post-adoption phases. The model also relates satisfaction and perceived usefulness to the degree to which the user's expectations about the service quality of the IS are confirmed. Expectations provide the baseline level against which confirmation is assessed by users to determine their satisfaction (Bhattacherjee 2001). The better they are met, the more useful it appears to users and the more satisfied they are. Consistent with prior research, we hypothesize that perceived usefulness will also play a significant role in SaaS user satisfaction and continuance intention.

H2: *Perceived usefulness has a positive association with SaaS continuance intention.*

H3: *Perceived usefulness has a positive association with satisfaction.*

In the marketing literature, much research effort has been directed at understanding the process of product/service evaluation involved in the formation of satisfaction response, considered key to acquiring and retaining loyal consumers (e.g. Oliver 1980). There is ample evidence that satisfaction is a function of perceived product/service performance and confirmation of performance expectations (Oliver and DeSarbo 1988). Perceived product/service performance is a direct result of evaluation of product/service attributes during the actual consumption experience. In general, one will be satisfied when the product/service performance is favorable and dissatisfied when the product/service performance is unfavorable. Confirmation of performance expectations is an evaluation of whether the performance of product/service is better than or worse than expected. The confirmation of expectations is described as the discrepancy or gap between prior expectations and actual performance of products/services and thus can be positive or negative. According to ECT, the more positive the confirmation of performance expectations, the greater the satisfaction (Yi 1990). Similarly, empirical studies in IS research have found that confirmation of performance expectations towards an IS tend to elevate users' perceived usefulness, while disconfirmation will reduce such perceptions (Bhattacherjee 2001; Limayem et al. 2007). Drawing on this research, we hypothesize that confirmation of SaaS service quality will also affect SaaS user satisfaction and perceived usefulness.

H4: *Confirmation of SaaS service quality has a positive association with satisfaction.*

H5: *Confirmation of SaaS service quality has a positive association with perceived usefulness.*

Research Methodology

Data Collection

To test the research model and validate our measurement instrument including a richer conceptualization of service quality, we designed questionnaires and conducted two surveys (i.e. one for pilot testing the SaaS service quality scale and one for testing our complete research model) with SaaS using firms based on data drawn from the Hoppenstedt firm database (Bisnode Business Information Group). To support the external validity of our study, we did not constrain both samples to specific industries or to firms of a specific organizational size. Data collected from the pilot test was used for instrument refinement and validation of factorial validity, whereas data from the second survey was used for confirmatory analysis of measurement properties and hypothesis testing of the research model. The survey questionnaires (see questionnaire items in Table 1 in the Appendix) for both samples were mailed to the senior-most IS manager in each SaaS-using firm (e.g. chief information officer [CIO], vice president in charge of IS), along with a letter outlining the purpose of the research, soliciting their participation in the survey, and a postage-paid return envelope for mailing back completed responses. To reduce self-reporting bias, each participant was given the opportunity to receive a report regarding how his/her firm position compares to firms of similar size and in similar industries. For the pilot test, we sent the questionnaire (that included the SaaS service quality questions only) to a random sample of 1,000 companies. After 41 responses were dropped due to missing data, a total of 111 usable responses from SaaS using firms could be used for our analysis. For the second survey, we sent a questionnaire with questions referring to all factors of our research model to another random sample of 2,000 companies. After sorting out 73 responses due to missing data, we received 172 usable responses from SaaS using firms for testing our research model.

The current study utilized a "key informants" methodology for data collection (Segars and Grover 1998). In survey research, targeted respondents assume the role of a key informant and provide information on a particular unit of analysis by reporting on group or organizational properties. However, if a respondent lacks appropriate knowledge, results can be confounding and may lead to erroneous conclusions. Therefore, within the context of this study, it was important to identify not only organizations that actively used SaaS applications, but also to identify respondents within those organizations who were intimately involved with, and most knowledgeable about the organization's SaaS usage. With this in mind, we introduced both of our surveys with a clear definition of SaaS services (in contrast to traditional IT outsourcing and ASP) and indicated that the survey should be filled out by the senior-most IS manager having a good overview of the organization's perception of SaaS service quality. Moreover, to increase the content validity of the answers, we asked the respondents to fill out the questionnaire regarding one specific SaaS application type (e.g., ERP or CRM) that they have in use.

To assess potential threats of nonresponse bias for both samples, the respondent and nonrespondent firms were compared with respect to sales, industry and the number of employees. No significant differences were found at the 0.05 level. For the sake of brevity, we present the demographic results only for our research sample (N=172) which were similar to the descriptives of our pilot sample. Demographic information about the respondents showed that about 47.1 percent were senior IS executives and 39.5 percent were IS managers. Although some preliminary steps were taken to ensure appropriate selection of key informants, a formal check was administered as part of the questionnaire (Kumar et al. 1993). Specifically, two items were used to assess an informant's length and frequency of SaaS usage (see Table 2). The mean score for the length of SaaS usage was 24.42 months and for the frequency of SaaS usage 18.56 times a month, indicating that respondents were appropriate and, thus, all responses were retained. More than 40 percent of the firms in the research sample have had SaaS in use for more than 3 years. 46.5 percent have been using SaaS between 1 and 3 years, and only 13.4 percent reported using SaaS less than a year. The research sample included firms with the following industry breakdown: manufacturing (29.07%), wholesale and retail trade (25.00%), financial intermediation (15.12%), TIME (telecommunication, information technology, media, entertainment) industries (11.05%), construction and real estate (8.72%), logistics (5.23%), public and healthcare (4.07%) and electricity/gas/water supply (1.74%). The sample split along enterprise application types were as follows: ERP (18.02%), SCM (14.53%), CRM (25.00%), CCC (Communication, Content, Collaboration) (19.19%), and Office (23.26%). Further sample characteristics are shown in Table 2.

Table 2. Sample Descriptives for Research Study (*N=172*)

Category	Percent	Category	Percent
Number of employees of SaaS using firms		*Annual sales (Euro million)*	
< 49	30.2	< 5	28.5
50 – 500	39.5	5 – 50	34.3
> 500	30.2	> 50	37.2
Number of years since SaaS rollout		*Respondent title*	
< 1	13.4	IS executives (CIO/CTO/VP of IS/IT)	47.1
1 – 3	46.5	Business executives (CEO, CFO and COO)	9.9
3 – 5	32.6	IS (middle) managers	39.5
> 5	7.6	Business managers and users	3.5
Length of SaaS usage of respondent		*Frequency of SaaS usage of respondent*	
< 3 months	6.3	4 or less times a month	5.7
3 to less than 6 months	7.7	5 to 8 times a month	15.3
6 to less than 12 months	17.6	9 to 12 times a month	33.9
12 months or more	68.4	13 or more times a month	45.1

Given the single method we used to collect data, we also conducted Harman's one-factor test (Podsakoff and Organ 1986). We performed an exploratory factor analysis on all the variables, but no single factor was observed and no single factor accounted for a majority of the covariance in the variables. Furthermore, a correlational marker technique was used, in which the highest variable from the factor analysis was entered as an additional independent variable (Richardson et al. 2009). This variable did not create a significant change in the variance explained in the dependent variables. Both tests suggest that common-method bias does not significantly impact our analyses and results. The data was thus deemed suitable for testing our hypothesized research model.

Measurement of Constructs and Instrument Validation

Scale refinement process for SaaS-QUAL scale

Searching for an appropriate measure for SaaS service quality (henceforth referred to as SaaS-QUAL), we thoroughly reviewed prior (IS) SERVQUAL measurements. As discussed above, we found several scales as points of departures for a preliminary scale (ASP-QUAL, E-S-QUAL etc.). Though these scales provided a solid conceptual basis for our study, several features of SaaS have not been reflected by them (e.g. granularity/modularity of software packages, interoperability, specific security/privacy aspects, traffic bottlenecks etc.). Therefore, we refined the SaaS-QUAL scale in a sequence of steps consistent with conventional guidelines for scale development

(Churchill 1979; Gerbing and Anderson 1988). This approach has been widely adopted by IS researchers and has worked well in producing measures with desirable psychometric properties. Figure 2 summarizes the steps of developing the SaaS-QUAL scale. While steps 3 and 4 were conducted based on the pilot sample, steps 5 to 7 were conducted with data of the research sample collected through the second survey (see results in the next section). Following this approach, we generated a purified list of 42 items reflecting six facets of SaaS service quality: *Rapport, Responsiveness, Reliability, Flexibility, Features, and Security/Privacy* (see the results of our exploratory and confirmatory factor analyses in Table 2 in the Appendix; our initial list of 89 items is omitted for brevity).

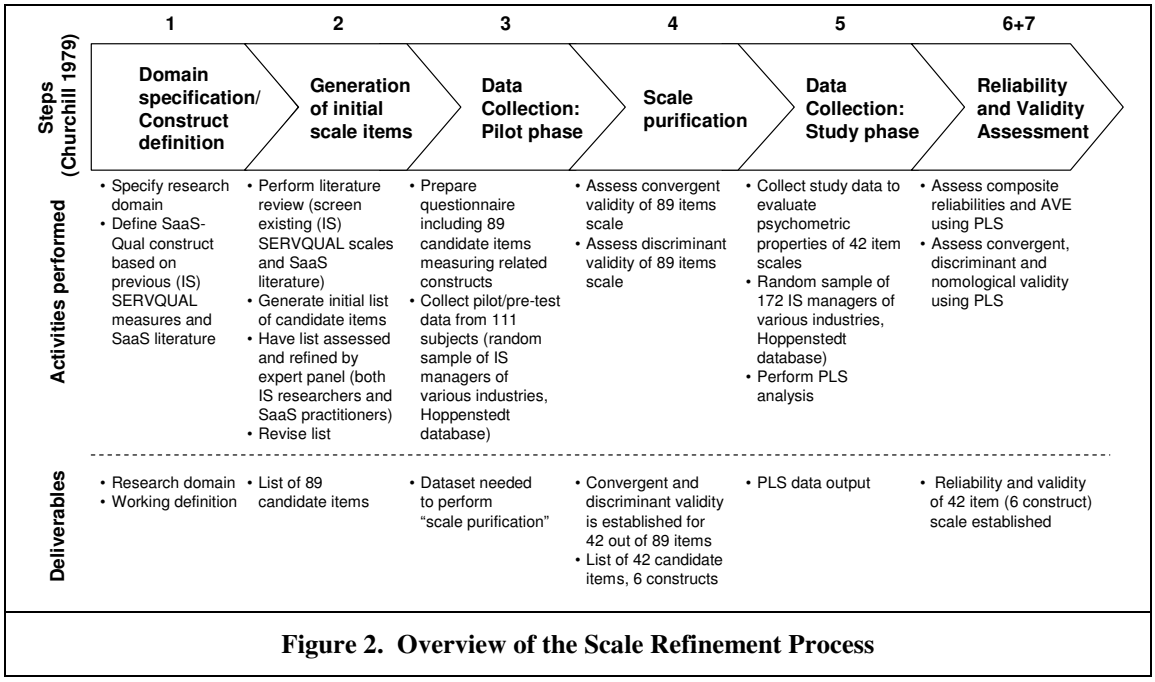

Figure 2. Overview of the Scale Refinement Process

Similar to Kettinger & Lee (2005), the dimensions 'Assurance' and 'Empathy' merged into one factor that we called 'Rapport'. Reliability and Responsiveness were validated as two separate constructs. Similar to Ma et al. (2005), 'Tangibles' was renamed to 'Features'. Two new constructs, i.e. 'Security/Privacy' and 'Flexibility', could be established for the SaaS-QUAL scale. Reflecting the content domain of the six constructs, we defined them as follows:

- '*Rapport*' includes all aspects of a SaaS provider's ability to provide knowledgeable, caring, and courteous support (e.g. joint problem solving or aligned working styles) as well as individualized attention (e.g. customer-specific trainings and courses).
- '*Responsiveness*' consists of all aspects of a SaaS provider's ability to ensure that the availability and performance of the SaaS-delivered application (e.g. through professional disaster recovery planning or load balancing) as well as the responsiveness of support staff (e.g. 24x7 hotline support availability) is guaranteed.
- '*Reliability*' comprises all features of a SaaS vendor's ability to perform the promised services dependably and accurately (e.g. providing services at the promised time).
- '*Flexibility*' refers to the degrees of freedom customers have to change contractual (e.g. cancellation period or payment model) or technical (e.g., scalability of storage capacity, individual modifications to the application service) aspects in the relationship with a SaaS vendor.
- '*Features*' refers to the degree the key features and functionalities (such as data extraction or application type specific functionality) of a SaaS application meet the business requirements of a customer.
- Finally, '*Security/Privacy*' includes all aspects to ensure that regular (preventive) measures (e.g. regular security audits, usage of encryption or anti-virus technology) are taken to avoid data breaches or system outages.

The psychometric properties of the six facets of SaaS service quality were further tested in our research sample which is reported in the next section.

Construct measurement and validation

The four constructs of interest in our research model were SaaS service quality (SaaS-Qual), satisfaction, perceived usefulness, and SaaS continuance intention. For each construct, we adopted validated measurement items from previous research studies (i.e. for satisfaction, perceived usefulness, and SaaS continuance intention) or developed them based on our own scale refinement procedures (i.e. for the SaaS-Qual scale), with minor changes on wording (see Table 1 in the Appendix). More specifically, satisfaction and SaaS continuance intention were measured as reflective first-order constructs with indicators adapted from Bhattacherjee (2001). Perceived usefulness was measured as a reflective first-order construct adapted from Davis et al. (1989).

SaaS service quality was measured as formative second-order construct with six reflective first-order constructs developed and validated in our scale development process (Ma et al. 2005; Parasuraman et al. 2005; Susarla et al. 2003). Our decision to use the reflective indicator specification for the first-order latent constructs is consistent with several key criteria recommended by Jarvis et al. (2003) for choosing that specification over the formative indicator specification (Jarvis et al. 2003): the relative homogeneity and hence interchangeability of scale items within each dimension, the high degree of covariation among items within each dimension, and the expectation that indicators within each dimension (e.g., Features) are likely to be affected by the same antecedents (e.g., Website design characteristics) and have similar consequences (e.g., increase or decrease in transaction speed or usability).

Since despite the reported lack of predictive validity, researchers have emphasized the diagnostic value and prescriptive validity of difference scores in the measurement of service quality factors, we adopted Kettinger & Lee's (2005) Zones of Tolerance (ZOT) approach that overcomes some of the most significant points of criticisms of the original SERVQUAL instrument. The ZOT approach measures service quality expectations on two levels including an adequate (minimum accepted) and a desired service level with their difference building the so-called 'Zone of Tolerance'. The approach suggests that the actual service performance should meet at least the minimum accepted service quality level to confirm customer expectations and yield satisfactory outcomes. We therefore used the difference scores between the actual service performance and the minimum accepted service quality level as measurement items for the confirmation of SaaS service quality. All other constructs were measured using seven-point Likert scales anchored with "strongly disagree" and "strongly agree".

Table 3. Assessment of Reflective Measurement Models: Loadings and Reliability

Constructs	*# of indicators*	*Range of Loadings*[1]	*Composite Reliability*	*AVE*
SaaS continuance intention	3	0.84 – 0.87	0.89	0.73
Satisfaction	4	0.94 – 0.97	0.98	0.92
Perceived usefulness	4	0.87 – 0.92	0.94	0.81
Rapport (P-M)[2]	9	0.79 – 0.96	0.87	0.83
Responsiveness (P-M)	9	0.86 – 0.97	0.81	0.87
Reliability (P-M)	5	0.81 – 0.96	0.86	0.83
Flexibility (P-M)	6	0.87 – 0.96	0.91	0.86
Features (P-M)	7	0.94 – 0.97	0.89	0.92
Security/Privacy (P-M)	6	0.70 – 0.95	0.85	0.77

[1] All factor loadings are significant at least at the $p<0.05$ level;
[2] (P-M) Difference score between perceived and minimum accepted service quality level

We assessed the psychometric properties of the measurement model results, seen in Table 3, by examining individual item loadings, internal consistency, convergent validity, and discriminant validity. The loadings of the measurement items on their respective factors were significant and above the threshold value of 0.70. Furthermore, measurement items did not have cross loadings above 0.4 on the unintended constructs (Hair et al. 2006) suggesting discriminant validity (the results of an principal components factor analysis are omitted here for brevity). The internal consistency of all reflective constructs clearly exceeded the threshold of 0.70, suggesting acceptable reliability. Convergent validity is considered adequate when the average variance extracted (AVE) is 0.50 or more. In addition, for satisfactory discriminant validity, the square root of average variance extracted (AVE) from the

construct should be greater than the variance shared between the construct and other constructs in the model. Table 4 lists the correlation matrix, with correlations among constructs and the square root of AVE on the diagonal. This table also provides strong evidence of discriminant validity.

Although the internal and external validity of the scales are typically assessed, there is a significant difference in the interpretation of the measurement model for formative constructs (Petter et al. 2007). Measures of internal consistency and unidimensionality cannot be used to judge the quality of the measurement model involving emergent constructs (Diamantopoulos and Winklhofer 2001). Accordingly, the general practice is to examine item weights. To estimate the formative second-order model of SaaS-Qual, we thus modeled the coefficients of each first-order factor to the second-order factor using a principal components factor analysis, following the procedure in Diamantopoulos and Winklhofer (2001). The assessment of SaaS-QUAL as second-order factor involved examining the correlations among the first-order factors. Tanaka and Huba (1984) argue for the possible validity of a second-order factor, if the first-order factors are highly correlated (Tanaka and Huba 1984). Table 4 shows that the first-order service quality factors are correlated and significantly different from zero, suggesting a second-order factor structure and validating their expected relationships.

Table 4. Correlation Matrix

Latent construct	1	2	3	4	5	6	7	8	9
(1) SaaS continuance int.	**0.85**								
(2) Satisfaction	0.50***	**0.96**							
(3) Perceived usefulness	0.60***	0.26**	**0.90**						
(4) Rapport (P-M)	0.33**	0.56***	0.21*	**0.91**					
(5) Responsiveness (P-M)	0.35**	0.66***	0.19*	0.54***	**0.93**				
(6) Reliability (P-M)	0.20*	0.54***	0.19*	0.51***	0.59***	**0.91**			
(7) Flexibility (P-M)	0.22*	0.62***	0.17*	0.49**	0.51***	0.42**	**0.93**		
(8) Features (P-M)	0.12ns	0.36**	0.25**	0.42**	0.43**	0.43**	0.41**	**0.96**	
(9) Security/Privacy (P-M)	0.16ns	0.64***	0.26**	0.47**	0.43**	0.49***	0.48***	0.41**	**0.88**

Note: Bolded diagonal elements are the square root of average variance extracted (AVE). These values should exceed inter-construct correlations (off-diagonal elements) for adequate discriminant validity. *p<0.05; **p<0.01; ***p<0.001; ns=not sign.

Because the correlations between all of the SaaS-QUAL facets are not negative, a high value on one factor does not preclude a high value on another. Moreover, the correlation among the first-order constructs are below the suggested cutoff value of 0.90 (Bagozzi et al. 1991), demonstrating that the content captured by the first-order factors are distinct from one another and indicative of discriminant validity. The coefficients (β-values) of the first-order enabling factors to the second-order factors are statistically significant (see Figure 3), providing justification for the existence of the hypothesized formative second-order model (Chin et al. 2003; Edwards 2001).

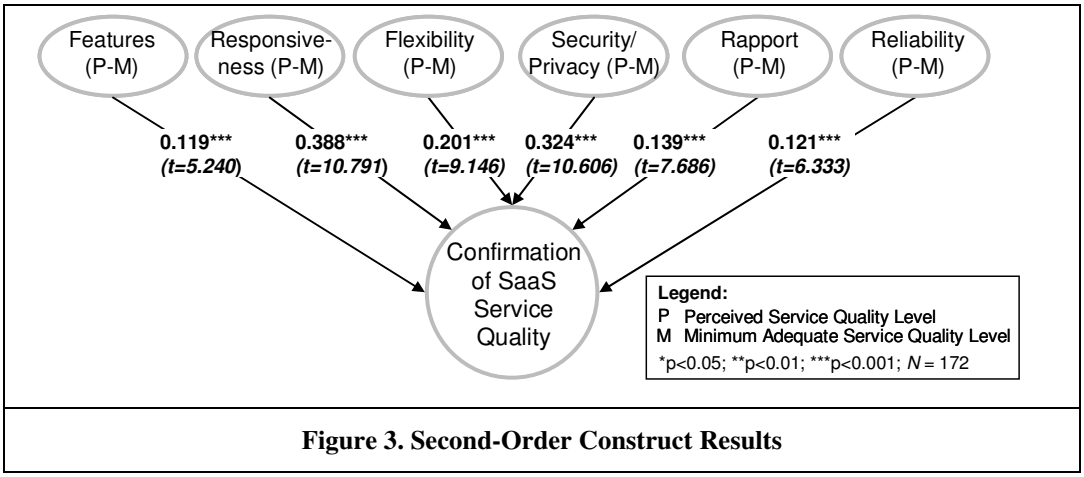

Figure 3. Second-Order Construct Results

Following the mediation tests using PLS, we also tested whether the second-order construct of overall SaaS-QUAL fully mediated the impact of the first-order facets on customer satisfaction. This step ensures that the second-order construct is a more parsimonious representation of the first-order constructs and fully captures their predictive power on the dependent variable it is theorized to predict. Overall, SaaS-QUAL was significant when all of the first-order factors were controlled suggesting it fully mediated the link between first-order constructs and customer satisfaction. This supports the conceptualization of SaaS-QUAL as second-order construct. Although PLS is reasonably robust against multicollinearity and skewed responses (Cassel et al. 1999), nonetheless, we performed the relevant assessment. We did not observe a high level of association between the exogenous constructs. Further, an examination of the variance inflation factors that ranged between 1.89 and 2.54 did not provide evidence of multicollinearity (Hair et al. 2006). Similarly, no evidence of heteroskedasticity was detected.

Data Analysis and Results

Descriptive Statistics: Meeting or Missing SaaS Customers' Zone of Tolerance?

In a first step, we analyzed the mean values of each of the six factors regarding their perceived service quality level and how they were positioned relative to the Zone of Tolerance (ZOT) (see Figure 4). All factors are meeting the ZOT except for two: *Responsiveness* ($\mu_{minimum}$=5.22; $\mu_{desired}$=6.49; $\mu_{performance}$=4.24) and *Security/Privacy* (μ_{min}=5.29; μ_{des}=6.49; μ_{per}=4.22) are far below the minimum acceptable service quality level. Notably, these are also the two factors with the highest values for minimum acceptable expectations and also the smallest ZOTs. While the perceived performance of *Reliability* (μ_{min}=3.73; μ_{des}=5.53; μ_{per}=4.43) and *Flexibility* (μ_{min}=4.03; μ_{des}=5.77; μ_{per}=4.26) was at the lower end of their ZOTs, *Rapport* (μ_{min}=2.69; μ_{des}=4.56; μ_{per}=4.15) was at the upper end. Lastly, the performance evaluations of *Features* (μ_{min}=3.43; μ_{des}=5.18; μ_{per}=4.37) hit the ZOT right at its center.

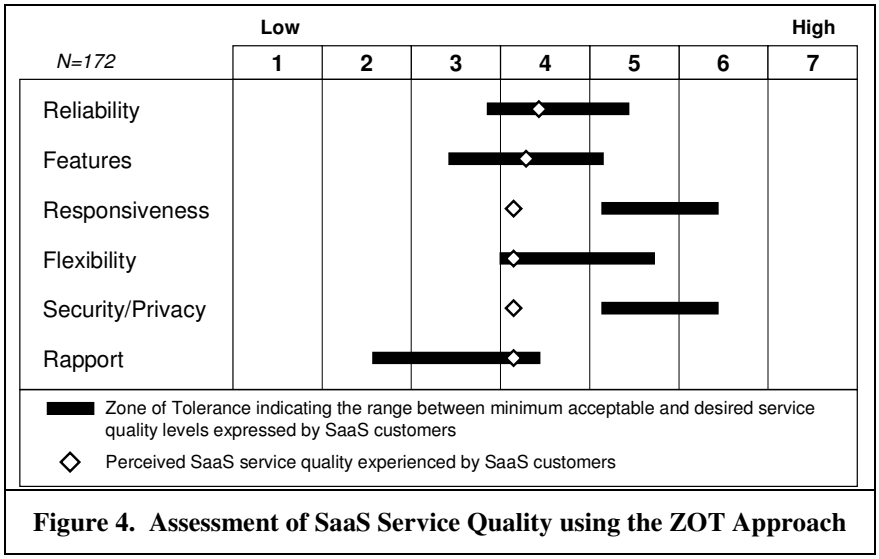

Figure 4. Assessment of SaaS Service Quality using the ZOT Approach

Structural Model Results

SmartPLS, version 2.0 (Ringle et al. 2005), was used to test the relationships among the study variables. Overall, our research model with the enriched conceptualization of SaaS service quality was supported (see Figure 5). First, the coefficients are in the appropriate direction and all are statistically significant. Second, the model explains a considerable portion of the variance in SaaS continuance intention (R^2=0.487), satisfaction (R^2=0.561), and perceived usefulness (R^2=0.263). These results attest to SaaS-QUAL's predictive validity. In an alternative structural model, we also tested the direct link between SaaS service quality confirmation and SaaS continuance intention. Consistent with previous studies, confirmation of SaaS service quality did not have a significant effect on SaaS continuance intention suggesting that its effect is fully mediated by perceived usefulness and satisfaction. Third, *Responsiveness* (β=.388; p<0.001) and *Security/Privacy* (β=.324; p<0.001) are the strongest factors contributing to SaaS service quality's impact on satisfaction and perceived usefulness.

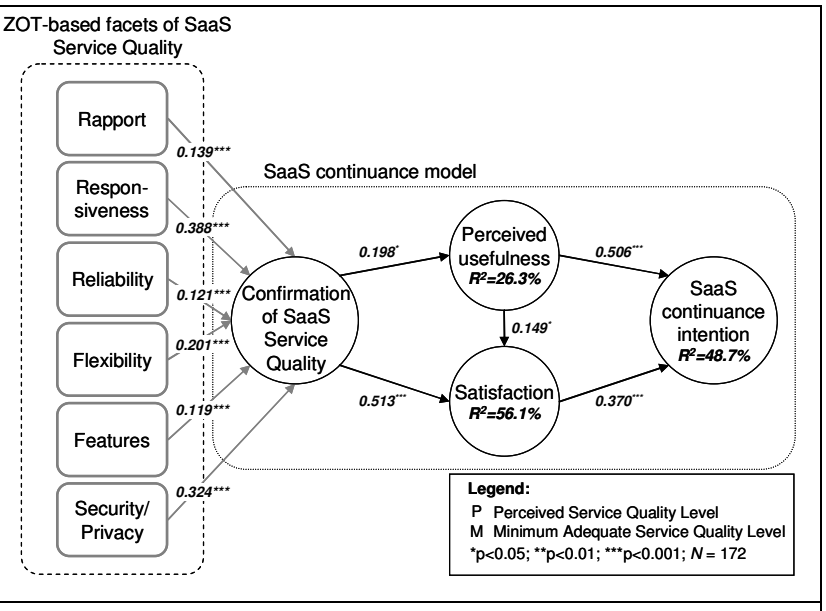

Figure 5. Structural Model Results

Discussion

Our study drew on previous service quality and SaaS literature to develop a service quality measure specifically for SaaS using companies. We refined the traditional SERVQUAL instrument and adopted the Zone of Tolerance (ZOT) approach. We then tested the instrument validity as well as its role in a nomological framework based on the IS continuance studies of Bhattacherjee (2001). Our results have provided several important theoretical insights as well as practical and methodological implications.

First, we found that confirmation of SaaS service quality is a significant driver of key antecedents of SaaS continuance intentions and has a particularly strong impact on customer satisfaction. In line with previous studies for other research contexts, SaaS service quality is fully mediated by perceived usefulness and, in particular, by customer satisfaction and thus only indirectly influences SaaS continuance intentions. Based on Bhattacherjee's model on IS continuance intentions, we zeroed in on the confirmation of SaaS service quality to understand in much more depth the role of SaaS service quality facets in shaping customer satisfaction, perceived usefulness and continued SaaS usage. As key SaaS-QUAL factors driving the influence on customer satisfaction and perceived usefulness, we found that *Responsiveness* and *Security/Privacy* have the strongest impact. Overall, with this more fine-grained conceptualization of service quality confirmation, the variance explained of customer satisfaction could be increased to 0.56, almost double of the variance explained in Bhattacherjee's (2001) initial model (0.33) and other follow-up models (e.g. Limayem et al. (2007) could explain 27% in the variance of customer satisfaction). Further, the variance explained of the other two dependent variables could also be increased considerably.

Second, we found that all factors were meeting their ZOTs except for those that had the strongest influence on satisfaction and perceived usefulness, the highest values for minimum expectations, and also the narrowest ZOTs. *Responsiveness* and *Security/Privacy* were far below the minimum acceptable service quality levels. From this picture it is clear to see where SaaS providers would have to start in an attempt to increase customer satisfaction, perceived usefulness and, indirectly, SaaS continuance intentions. Since *Responsiveness* and *Security/Privacy* turn out to undercut their respective ZOTs, responsiveness and security/privacy-related issues should be addressed first. Since *Reliability* and *Flexibility* are at the bottom end of their ZOTs, there is still considerable room for improvement here as well. Last and with least weight, *Rapport* and *Features* show very positive values. SaaS providers thus do not need to worry about the feature spectrum of their SaaS-based applications nor about the relationship management skills of their service workers in the short and medium term.

Third, our results underscore the need for SaaS vendors and customers to place extra emphasis on attributes pertaining to *Responsiveness* and *Security/Privacy*. In this regard, it is noteworthy that in contrast to E-Commerce websites where consumers consider ease and speed of using the site to be the most important facets and privacy and

system availability the least important facets of Web Site quality (Parasuraman et al. 2005), the most crucial factors of SaaS quality service are pertaining to the behind-the-scenes infrastructure. Thus, earning a high-quality image for a SaaS solution involves much more than creating an excellent façade for a website. The perceptual attributes that constitute *Responsiveness* and *Security/Privacy* suggest that companies may not have full control over performance on this dimension. Internet network bandwidth and hacker attacks are for example also likely to affect the performance. SaaS vendors should thus be (a) sensitive to potential deleterious effects of system performance and availability and (b) proactive in identifying aspects of *Responsiveness* and *Security/Privacy* that are beyond their control (e.g. seeking to cover and provide a seamless and comprehensive service value chain from service origination to service delivery at customers' user PCs) and (c) devising appropriate communication scripts to appease complaining customers. Further, previous research in E-Commerce has argued that security and privacy of Web sites may not be critical for more frequent users (Wolfinbarger and Gilly 2003). Experience may also mitigate concerns about SaaS privacy and security. However, the fact that the respondents in our survey had sufficient experience with SaaS usage, coupled with the consistent findings that security/privacy perceptions do significantly influence customers' overall quality/value perceptions, emphasizes the need for SaaS vendors to continue to reassure customers through Web site design cues and external communications signaling the privacy/security of their sites.

Fourth, IT managers responsible for selecting or renewing SaaS-based solutions may learn from this study what a representative sample of IS executives considers to be the most important service quality factors in SaaS and where SaaS providers have to improve to meet customer satisfaction and perceived usefulness. As discussed above, IT managers should particularly focus on a SaaS vendor's operations management capabilities in the areas *Responsiveness* and *Security/Privacy*. More specifically, they could negotiate contractual uptime guarantees and IT helpdesk/application response time including penalties and escalation clauses if the performance standards are not achieved. On the security side, companies should place particular importance on defining careful and granular SLAs on security/privacy aspects including clear data protection and backup policies and regular audits of SLA compliance. By making potentially hidden expectations transparent, the regular tracking of SaaS-QUAL results may also be used to (further) inform and specify contractual elements of SLAs such as service level contents (e.g. targets, time frame), plans for future demand and change management (e.g. joint demand forecasting process), communication procedures (e.g. communication schedules and format), measurement charters (e.g. KPI metrics) and enforcement plans (e.g. penalty/reward definitions).

Finally, we refined, tested and validated an adapted SaaS-QUAL scale and thus add to existing SERVQUAL research by adapting it to a specific service context. It includes service factors that are conceptually different to other software service settings. In contrast to the ASP service quality scales developed so far, we found that the SaaS-QUAL scale comprises a factor called *Responsiveness* that conceptually enriches existing 'availability' factors found in the development of ASP-QUAL scales. Also, we identified a construct called *Security/Privacy* that conceptually integrates companies' security and privacy concerns extending existing security factors in ASP service quality scales. Finally, due to the specific nature of the SaaS business model, we could also determine an additional factor called *Flexibility* that has not been identified in other software service settings including ASP and E-Commerce. The SaaS-QUAL scale with its six dimensions demonstrated good psychometric properties and turned out to provide high nomological validity in the context of SaaS continuance intentions. Used as a diagnostic instrument, it may help SaaS providers spotting weak points in their service provisioning and thus direct investments accordingly (e.g. into better security measurements or trust-building activities). Trends in the dimensional- and attribute-level ratings from such tracking studies will further help identify software services' strengths and weaknesses and thus suggest ideas for improvement.

Limitations

Our research study involves several limitations. First, our findings must be interpreted in light of the limitations of cross-sectional research. Because our data are not longitudinal, we are unable to conclusively confirm the direction of causality. While we feel that the balance of logic in our study supports the idea that meeting service quality standards play an important role in the formation of satisfaction with and perceived usefulness of SaaS usage, longitudinal (i.e. multi-stage) research would help researchers to better understand the temporal relationships (e.g. how satisfaction in an earlier stage influences expectation levels in later stages) between our constructs (Watson et al. 1998). Second, we collected data from a single respondent within the organization. Given the nature of the survey items, the majority of respondents are IS senior executives with comprehensive understanding of the organization-wide SaaS usage. The respondent characteristics suggest good data quality, minimizing the potential problem of

single respondent bias. The tests conducted on our data also indicate that common method bias may not significantly affect our results. Nonetheless, there still exist concerns with analyses based on self-reported data collected from a single source. Third, the analysis of relevant sub-samples including different industries, firm sizes, or user groups would contribute to identifying the boundary conditions for the theoretical model and would add additional practical insights into the most important drivers of SaaS service quality across these sub-samples. Finally, since it was one of our explicit goals to develop and validate a highly practical measurement instrument, we used a ZOT-based difference score approach to measure the confirmation of SaaS service quality. Some researchers found in previous studies that the difference score method is vulnerable to reliability, validity, dimensionality and interpretability issues and should be replaced by a single item and perceptions-only measurement (e.g. SERVPERF) that demonstrates superior predictive and convergent validity. Others have emphasized the diagnostic and thus practical value of the difference score-based (IS) service quality measurements. Though our study demonstrated satisfying predictive values of SaaS service quality, future studies that may want to use a more parsimonious measurement instrument could replicate this study with alternative measurement models (e.g. with perceptions-only items or perception and exception items as separate component scores) and compare the predictive values of the different SaaS-QUAL scales.

Conclusion

Based on survey results from SaaS using firms of various industries, we have derived some important insights into how SaaS service quality factors affect IS continuance intentions by virtue of influencing customer satisfaction and perceived usefulness and how these factors are currently fulfilled from a (business) customer perspective. Specifically, the survey data show that the most important service quality factors in SaaS are *Responsiveness* and *Security/Privacy*. Interestingly, SaaS providers are still struggling to fulfill exactly these two factors. These insights may help as a starting point for SLA refinements between SaaS providers and customers and may well serve for resource allocation improvements. Moreover, regularly tracking customers' zones of tolerance may be used as a leading indicator for typical key performance factors such as monthly recurring revenue, churn/renewal rate, and cash flow.

As in the face of a growing service orientation in the IS industry as well as in IS research, SaaS-based software delivery is likely to gain importance, regularly assessing the service quality factors of SaaS services and their importance for continued SaaS usage will become even more critical for SaaS vendors and user organizations. Based on this study's findings, future research is encouraged to further examine the role of SaaS service quality in the IS post-adoption context. Particularly, studies analyzing how the impact of SaaS service quality on customer satisfaction and perceived usefulness changes along different post-adoption phases (i.e. acceptance, routinization, infusion) should yield valuable insights for organizations' continued SaaS usage behavior and their management of SaaS providers.

References

Bagozzi, R.P., Y., Y., and Phillips, L.W. 1991. "Assessing Construct Validity in Organizational Research," *Administrative Science Quarterly* (36:3), pp 421-458.

Bajaj, A., and Nidumolu, S.R. 1998. "A Feedback Model to Understand Information System Usage," *Information & Management* (33:4), pp 213-224.

Benlian, A. 2009. "A Transaction Cost Theoretical Analysis of Software-as-a-Service (Saas)-Based Sourcing in Smbs and Enterprises," *Proceedings of the 17th European Conference on Information Systems*, Verona, Italy.

Bhattacherjee, A. 2001. "Understanding Information Systems Continuance: An Expectation-Confirmation Model," *MIS Quarterly* (25:3), pp 351-370.

Brown, T.J., Churchill Jr, G.A., and Peter, J.P. 1993. "Improving the Measurement of Service Quality," *Journal of Retailing* (69:1), Spring93, p 127.

Cassel, C., Hackl, P., and Westlund, A.H. 1999. "Robustness of Partial Least-Squares Method for Estimating Latent Variable Quality Structures," *Journal of Applied Statistics* (26:4), pp 435-446.

Chin, W.W., Marcolin, B.L., and Newsted, P.R. 2003. "A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and an Electronic-Mail Emotion/Adoption Study," *Information Systems Research* (14:2), pp 189-217.

Choudhary, V. 2007. "Comparison of Software Quality under Perpetual Licensing and Software as a Service," *Journal of Management Information Systems* (24:2), pp 141-165.

Churchill, G.A. 1979. "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research* (16:1), pp 64-73.

Compeau, D., Higgins, C.A., and Huff, S. 1999. "Social Cognitive Theory and Individual Reactions to Computing Technology: A Longitudinal Study," *MIS Quarterly* (23:2), pp 145-158.

Cronin Jr, J.J., and Taylor, S.A. 1992. "Measuring Service Quality: A Reexamination and Extension," *Journal of Marketing* (56:3), pp 55-68.

Davis, F.D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly* (13:3), pp 319-339.

Davis, F.D., Bagozzi, R.P., and Warshaw, P.R. 1989. "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," in: *Management Science*. INFORMS: Institute for Operations Research, pp. 982-1003.

Deng, L., Turner, D.E., Gehling, R., and Prince, B. 2010. "User Experience, Satisfaction, and Continual Usage Intention of It," *European Journal of Information Systems* (19:1), pp 60-75.

Diamantopoulos, A., and Winklhofer, H.M. 2001. "Index Construction with Formative Indicators: An Alternative to Scale Development," *Journal of Marketing Research* (38:2), pp 269-277.

Edwards, J.R. 2001. "Multidimensional Constructs in Organizational Behavior Research: An Integrative Analytical Framework," *Organizational Research Methods* (4:2), pp 144-192.

Gerbing, D.W., and Anderson, J.C. 1988. "An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment," *Journal of Marketing Research* (25:2), pp 186-192.

Hair, J.F., Anderson, R.E., Tatham, R.L., and Black, W.C. 2006. *Multivariate Data Analysis with Readings*. Englewood Cliffs, N.J.: Prentice Hall.

Hsu, M., Chiu, C., and Ju, T. 2004. "Determinants of Continued Use of the Www: An Integration of Two Theoretical Models," *Industrial Management & Data Systems* (104:9), pp 766-775.

Jarvis, C.B., Mackenzie, S.B., Podsakoff, P.M., Mick, D.G., and Bearden, W.O. 2003. "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research," *Journal of Consumer Research* (30:2), pp 199-218.

Jasperson, J., Carter, P.E., and Zmud, R.W. 2005. "A Comprehensive Conceptualization of Post-Adoptive Behaviors Associated with Information Technology Enabled Work Systems," *MIS Quarterly* (29:3), pp 525-557.

Jiang, J.J., Klein, G., and Carr, C.L. 2002. "Measuring Information System Service Quality: Servqual from the Other Side," *MIS Quarterly* (26:2), pp 145-166.

Karahanna, E., Straub, D.W., and Chervany, N.L. 1999. "Information Technology Adoption across Time: A Cross-Sectional Comparison of Pre-Adoption and Post-Adoption Beliefs," *MIS Quartely* (23:2), pp 183-213.

Kettinger, W.J., and Lee, C.C. 1997. "Pragmatic Perspectives on the Measurement of Information Systems Service Quality," *MIS Quarterly* (21:2), pp 223-240.

Kettinger, W.J., and Lee, C.C. 2005. "Zones of Tolerance: Alternative Scales for Measuring Information Systems Service Quality," *MIS Quarterly* (29:4), pp 607-623.

Kohlmeyer, J.M., and Blanton, J.E. 2000. "Improving Is Service Quality," *Journal of Information Theory and Applications* (2:1), pp 1-10.

Kumar, N., Stern, L.W., and Anderson, J.C. 1993. "Conducting Interorganizational Research Using Key Informants," *Academy of Management Journal* (36:6), pp 1633-1651.

Lewis, R.C., and Booms, B.H. 1983. "The Marketing Aspects of Service Quality," in: *Emerging Perspectives on Services Marketing,* L.L. Berry, G.L. Shostack and G. Upah (eds.). Chicago: American Marketing Association, pp. 99-107.

Limayem, M., Hirt, S.G., and Cheung, C.M.K. 2007. "How Habit Limits the Predictive Power of Intention: The Case of Information Systems Continuance," *MIS Quarterly* (31:4), pp 705-737.

Ma, Q., Pearson, J.M., and Tadisina, S. 2005. "An Exploratory Study into Factors of Service Quality for Application Service Providers," *Information & Management* (42:8), pp 1067-1080.

Menken, I. 2008. *Saas - the Complete Cornerstone Guide to Software as a Service Best Practices: Concepts, Terms, and Techniques for Successfully Planning, Implementing and Managing Saas Solutions*. Emereo.

Mertz, S.A., Eschinger, C., Eid, T., Huang, H.H., Pang, C., and Pring, B. 2009. "Market Trends: Software as a Service, Worldwide, 2008-2013," Gartner.

Oliver, R.L. 1980. "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research (JMR)* (17:4), pp 460-469.
Oliver, R.L., and DeSarbo, W.S. 1988. "Response Determinants in Satisfaction Judgments," *Journal of Consumer Research* (14:4), pp 495-507.
Parasuraman, A., Berry, L.L., and Zeithaml, V.A. 1991. "Refinement and Reassessment of the Servqual Scale," *Journal of Retailing* (67:4), Winter91, p 420.
Parasuraman, A., and Zeithaml, V.A. 2002. "Measuring and Improving Service Quality: A Literature Review and Research Agenda," in: *Handbook of Marketing,* B. Weitz (ed.). Thousand Oaks, CA: Sage.
Parasuraman, A., Zeithaml, V.A., and Berry, L.L. 1985. "A Conceptual Model of Service Quality and Its Implications for Future Research," *Journal of Marketing* (49:4), Fall85, pp 41-50.
Parasuraman, A., Zeithaml, V.A., and Malhotra, A. 2005. "E-S-Qual: A Multiple-Item Scale for Assessing Electronic Service Quality," *Journal of Service Research* (7:3), February 1, 2005, pp 213-233.
Petter, S., Straub, D., and Rai, A. 2007. "Specifying Formative Constructs in Information Systems Research," *MIS Quarterly* (31:4), pp 623-656.
Pettey, C. 2006. "Gartner Says 25 Percent of New Business Software Will Be Delivered as Software as a Service by 2011." Stamford, Conn.: Gartner.
Pitt, L.F., Watson, R.T., and Kavan, C.B. 1997. "Measuring Information Systems Service Quality: Concerns for a Complete Canvas," *MIS Quarterly* (21:2), pp 209-221.
Podsakoff, P.M., and Organ, D.W. 1986. "Self-Reports in Organizational Research: Problems and Prospects," *Journal of Management* (12:4), Winter86, p 531.
Pring, B., and Lo, T. 2009. "Dataquest Insight: Saas Adoption Trends in the U.S. And U.K.," Gartner.
Rai, A., and Sambamurthy, V. 2006. "The Growth of Interest in Services Management: Opportunities for Information Systems Scholars," *Information Systems Research* (17:4), pp 327-331.
Richardson, H.A., Simmering, M.J., and Sturman, M.C. 2009. "A Tale of Three Perspectives: Examining Post Hoc Statistical Techniques for Detection and Correction of Common Method Variance," *Organizational Research Methods* (12:4), pp 762-800.
Ringle, C.M., Wende, S., and Will, A. 2005. "Smartpls 2.0 (M3) Beta," University of Hamburg, Hamburg, http://www.smartpls.de.
Segars, A., and Grover, V. 1998. "Strategic Information Systems Planning Success: An Investigation of the Construct and Its Measurement," *MIS Quartely* (22:2), pp 139-163.
Sigala, M. 2004. "The Asp-Qual Model: Measuring Asp Service Quality in Greece," *Managing Service Quality* (14:1), pp 103-114.
Susarla, A., Barua, A., and Whinston, A. 2003. "Understanding the Service Component of Application Service Provision: An Empirical Analysis of Satisfaction with Asp Services," *MIS Quarterly* (27:1), pp 91-123.
Tanaka, J.S., and Huba, G.J. 1984. "Confirmatory Hierarchical Factor Analysis of Psychological Distress Measures," *Journal of Personality and Social Psychology* (46:3), pp 621-635.
Teas, R.K. 1993. "Expectations, Performance Evaluation, and Consumers' Perceptions of Quality," *Journal of Marketing* (57:1), pp 18-34.
van Dyke, T.P., Kappelman, L.A., and Prybutok, V.R. 1997. "Measuring Information Systems Service Quality: Concerns on the Use of the Servqual Questionnaire," *MIS Quarterly* (21:2), pp 195-208.
van Dyke, T.P., Prybutok, V.R., and Kappelman, L.A. 1999. "Cautions on the Use of the Servqual Measure to Assess the Quality of Information Systems Services," *Decision Sciences* (30:3), Summer99, pp 877-891.
Watson, R.T., Pitt, L.F., and Kavan, C.B. 1998. "Measuring Information Systems Service Quality: Lessons from Two Longitudinal Case Studies," *MIS Quarterly* (22:1), pp 61-79.
Weier, M.H. 2009. "Google, Workday Outages Show Saas Isn't Perfect," in: *Information Week.*
Wolfinbarger, M., and Gilly, M.C. 2003. "Etailq: Dimensionalizing, Measuring, and Predicting Etail Quality," *Journal of Retailing* (79:3), pp 183-198.
Xin, M., and Levina, N. 2008. "Software-as-a Service Model: Elaborating Client-Side Adoption Factors," *Proceedings of the Twenty Ninth International Conference on Information Systems, Paper 86*, Paris, France.
Yi, Y. 1990. " A Critical Review of Consumer Satisfaction," in: *Review of Marketing,* V.A. Zeithaml (ed.). Chicago, IL: American Marketing Association, pp. 68-123.
Zeithaml, V.A., Berry, L.L., and Parasuraman, A. 1993. "The Nature and Determinants of Customer Expectations of Service," *Journal of the Academy of Marketing Science* (21:1), pp 1-12.
Zeithaml, V.A., Parasuraman, A., and Malhotra, A. 2002. "Service Quality Delivery through Web Sites: A Critical Review of Extant Knowledge," *Journal of the Academy of Marketing Science* (30:4), pp 362-375.

Appendix

Table 1. Measurement Scales (Research Study, *N*=172)

Constructs	***Indicators***	
SaaS Continuance Intention (SCI) (Bhattacherjee 2001)	SCI1	We intend to continue using SaaS rather than discontinue its use.
	SCI2	Our intentions are to continue using SaaS than use any alternative means (on-premise solutions).
	SCI3	If I could, I would like to discontinue my use of SaaS (reverse coded).
Satisfaction (S) (Bhattacherjee 2001)	How do you feel about your overall experience of SaaS use?	
	S1	Very dissatisfied/Very satisfied.
	S2	Very displeased/Very pleased.
	S3	Very frustrated/Very contented.
	S4	Absolutely terrible/Absolutely delighted.
Perceived Usefulness (PU) (Davis et al. 1989)	PU1	Using SaaS improves our performance in managing our functions/processes.
	PU2	Using SaaS increases our productivity in managing our functions/processes.
	PU3	Using SaaS enhances our effectiveness in managing our functions/processes.
	PU4	Overall, SaaS is useful in supporting our functions/processes.

ZOT-based SaaS-QUAL's Anchor Questions and Format for Item Descriptions (see below)
(Source: Kettinger and Lee 2005; Ma et al. 2005; Parasuraman et al. 1991, 2005; Susarla et al. 2003; Sigala 2004; Zeithaml et al. 2002)

Facets of overall SaaS-QUAL	When it comes to …	My Minimum Service Level is: (1=Low, 7=High)	My Desired Level of Service is: (1=Low, 7=High)	My Perception of the Performance of the Service is: (1=Low, 7=High)
Rapport (Ra)	… a shared approach to problem-solving			
	… customer-specific trainings and courses			
	… understanding our business goals and processes			
	… a good personal relationship (e.g. SaaS provider's employees are consistently courteous with our users)			
	… an aligned working style (e.g. convenient operating hours)			
	… to consulting services			
	… the cultural fit between SaaS provider and our company			
	… support that is tailored to our individual needs (e.g. documentation materials, etc.)			
	… the SaaS provider's knowledge to do its job well			
Responsiveness (Res)	… system availability/uptime for business (e.g. system crash or freeze)			
	… network performance			
	… efficient disaster recovery			
	… efficient contingency and replacement policy			
	… hardware and software redundancy			
	… an adequate number of service personnel dedicated to our company			
	… the support of up-to-date, cutting-edge hardware, software and netware technology			
	… (technical) support availability (i.e. the promptness of providing services)			
	… (multi-channel) customer care (i.e. the SaaS provider's willingness to help users)			
Reliability (Rel)	… providing services at the promised time			
	… performing services right the first time			
	… contract fulfillment of the service provider (including services and the times it promises to do so)			
	… user problems, our SaaS provider shows a sincere interest in solving it (e.g. incident resolution time)			
	… the provision of error-free services and accurate budgetary controls			
Flexibility (Fl)	… the integration capability of our outsourced SaaS application with our ICT infrastructure			
	… application scalability			
	… modularity of features (i.e. packaging choices)			
	… application customization (i.e. configurability)			
	… modifying contractual parameters at later stages			
	… choices of ways to pay (e.g. payment/billing options)			
Features (Fe)	… a visually appealing and sympathetic user interface			
	… a user-friendly navigation structure and search functionality			
	… data reporting and extracting features			
	… the SaaS application's configuration (e.g. user administration etc.) features			
	… upgrading the SaaS application			
	… the SaaS application's dashboard features with metrics measuring customers service usage			
	… the SaaS application's core functionality to support process steps/activities			
Security/ Privacy (SePr)	… protection against data loss/corruption			
	… regular security audits			
	… providing a secure physical environment (secure data center)			
	… anti-virus protection			
	… data encryption			
	… data confidentiality (i.e. keeping data private)			

Table 2. CFA and EFA Results for the SaaS-QUAL Scale (Pilot study, *N*=111)

	Range of CFA loadings		*Range of EFA Loadings (after oblique rotation)*[a]					
Factors	***Loadings***[b]	***t-value***[c]	***Rapport***	***Responsiven.***	***Reliability***	***Flexibility***	***Features***	***Sec./Priv.***
Rapport (coefficient alpha = .89)								
Ra1	.82	22.01	.80					
Ra2	.80	18.34	.78					
Ra3	.79	17.54	.77					
Ra4	.77	16.43	.73					
Ra5	.79	18.10	.75					
Ra6	.82	23.72	.82					
Ra7	.87	25.82	.89					
Ra8	.84	24.24	.81					
Ra9	.82	22.33	.78					
Responsiveness (coefficient alpha = .83)								
Res1	.79	15.32		.73				
Res2	.77	13.73		.71				
Res3	.78	14.43		.72				
Res4	.81	17.27		.82				
Res5	.73	12.30		.74				
Res6	.79	16.03		.79				
Res7	.76	12.57		.70				
Res8	.80	17.03		.81				
Res9	.80	17.16		.75				
Reliability (coefficient alpha = .93)								
Rel1	.83	26.73			.83			
Rel2	.85	27.92			.80			
Rel3	.86	27.02			.94			
Rel4	.84	26.04			.82			
Rel5	.81	18.03			.78			
Flexibility (coefficient alpha = .86)								
Fl1	.79	14.72				.75		
Fl2	.78	13.94				.72		
Fl3	.82	19.93				.79		
Fl4	.80	20.32				.80		
Fl5	.81	20.19				.78		
Fl6	.76	18.76				.70		
Features (coefficient alpha = .94)								
Fe1	.85	21.03					.81	
Fe2	.86	21.77					.89	
Fe3	.84	20.89					.82	
Fe4	.87	22.15					.91	
Fe5	.85	21.37					.83	
Fe6	.84	20.63					.79	
Fe7	.85	20.89					.81	
Security/Privacy (coefficient alpha = .90)								
SePr1	.79	16.03						.78
SePr2	.79	15.93						.75
SePr3	.83	17.87						.77
SePr4	.85	19.92						.82
SePr5	.81	17.01						.87
SePr6	.84	18.84						.81
SePr7	.88	15.93						.94
Range of Goodness-of-fit statistics in CFA: X^2/df=2.56, CFI=0.97, NFI=0.96, RFI=0.96, TLI=0.96, RMSEA= 0.04								

NOTE: CFA = confirmatory factor analysis; EFA = exploratory factor analysis; CFI = Comparative Fit Index; NFI = Normed Fit Index; RFI = Relative Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation.
a. Total variance extracted by the six factors = 84%; rotation method: oblimin with Kaiser normalization; average interfactor correlation = .42
b. These are standardized loading estimates from CFA using the Amos software package.
c. Based on one-tailed tests, t-values greater than 1.65 are significant at p<.05; t-values greater than 2.33 are significant at p<.01.