Service Quality in Software-as-a-Service: Developing the SaaS-Qual Measure and Examining Its Role in Usage Continuance

ALEXANDER BENLIAN, MARIOS KOUFARIS, AND THOMAS HESS

ALEXANDER BENLIAN is an assistant professor of business administration and management information systems at the University of Munich, Germany. He holds a Ph.D. in business administration and management information systems from the University of Munich. His research interests include IT services, evaluation and selection of enterprise application systems, and recommender systems in online commerce. His work has been published in international journals such as *Information Systems Journal, European Journal of Information Systems, International Journal of Electronic Commerce, Decision Support Systems, Information & Management, Journal of Business & Information Systems Engineering,* and in conference proceedings, such as those of the International Conference on Information Systems (ICIS), European Conference on Information Systems (ECIS), and Americas Conference on Information Systems (AMCIS).

MARIOS KOUFARIS is an associate professor in computer information systems at the Zicklin School of Business of Baruch College, City University of New York. He received a Ph.D. in information systems from the Stern School of Business of New York University, as well as a B.Sc. in decision sciences from the Wharton School of Business and a B.A. in psychology from the College of Arts and Sciences, both at the University of Pennsylvania. His research interests include consumer behavior in online commerce, end-user behavior, and the social impact of IT. His work has been published in *Information Systems Research, Journal of Management Information Systems, MIS Quarterly, International Journal of Electronic Commerce, Information & Management, DATA BASE for Information Systems,* and *Communications of the ACM.*

THOMAS HESS is a professor of management information systems at the University of Munich, Germany, where he also serves as director of the Institute for Information Systems and New Media and as coordinator of the Center for Internet Research and Media Integration. He holds a Ph.D. in management from the University of St. Gallen, Switzerland. His research interests include digitalization strategies, digital products, media management, and economics of information systems. His work has appeared in international journals as *Communications of the ACM, Information Systems Journal, International Journal of Electronic Commerce, Decision Support Systems, Electronic Markets, Information & Management,* and *Information Society.* Professor Hess has also published in the proceedings of conferences such as International Conference on Information Systems (ICIS) European Conference on Information Systems (ECIS), and Americas Conference on Information Systems (AMCIS).

ABSTRACT: Despite the need to better understand how customers of software-as-a-service (SaaS) solutions perceive the quality of these software services and how these

Journal of Management Information Systems / Winter 2011–12, Vol. 28, No. 3, pp. 85–126.
© 2012 M.E. Sharpe, Inc.
0742–1222 / 2012 $9.50 + 0.00.
DOI 10.2753/MIS0742-1222280303

perceptions influence SaaS adoption and use, there is no extant measure that comprehensively captures service quality evaluations in SaaS. Based on previous SERVQUAL and SaaS literature, field interviews and focus groups, a card-sorting exercise, and two surveys of SaaS using companies, we develop, refine, and test SaaS-Qual, a zones-of-tolerance (ZOT)–based service quality measurement instrument specifically for SaaS solutions. Besides validating already established service quality dimensions (i.e., rapport, responsiveness, reliability, and features), we identify two new factors (i.e., security and flexibility) that are essential for the evaluation of service quality of SaaS solutions. SaaS-Qual demonstrates strong psychometric properties and shows high nomological validity within a framework that predicts the continued use of SaaS solutions by existing customers. In addition to developing a validated instrument that provides a fine-grained measurement of SaaS service quality, we also enrich existing research models on information systems continuance. Moreover, the SaaS-Qual instrument can be used as a diagnostic tool by SaaS providers and users alike to spot strengths and weaknesses in the service delivery of SaaS solutions.

Key words and phrases: IS continuance, SaaS-Qual, service quality, SERVQUAL, software-as-a-service, zones of tolerance.

Software-as-a-service (SaaS), which describes software applications delivered as a service over the Internet [2, 23], is quickly becoming an important model of software delivery for companies of all sizes and in all industries [83, 84]. For software users, SaaS is said to provide numerous benefits, including IT cost reductions, operational elasticity, faster upgrade cycles, and ease of implementation [2]. For software providers, SaaS is an increasingly significant channel for selling software services and challenges conventional business models of software firms. Worldwide software revenues for SaaS delivery are forecast to grow by 19.4 percent overall from 2008 to 2013, which is more than triple the total market compound annual growth rate of 5.2 percent [58]. This is especially true in those application markets where low levels of system customization are required (e.g., office suites, collaboration) [68].

However, in order for SaaS to grow beyond its initial diffusion stage, it must be perceived by its costumers as an effective and efficient alternative to traditional software models. This means that service quality issues are pivotal to its continued success. Recent reports have shown instances where the provision of SaaS offerings has missed customers' service quality expectations, and as SaaS becomes more prevalent, such instances are increasing [91]. A recent Gartner study of 333 organizations based in the United States and the United Kingdom [72] has illustrated that the top three reasons why organizations discontinue SaaS are unfulfilled technical requirements, security issues, and low-quality customer support. The failure to fulfill customers' expectations regarding service quality, such as application availability or vendor responsiveness, may thus have critical consequences not only for the customers but also for the vendors.

The emergence and advancement of SaaS services and the increasing pervasiveness of the cloud infrastructure that underlies SaaS services have introduced new requirements toward software service quality [2]. For example, in a software usage environment such as SaaS that is heavily dependent on the Internet infrastructure, data security and service availability become increasingly important [56]. The amount of flexibility afforded by the SaaS service is also a significant concern for SaaS customers. Since they do not have to own and maintain the infrastructure necessary to run the software and they pay for SaaS services using very flexible payment models (e.g., subscription or pay-as-you-go models), they can switch SaaS vendors more easily, leading to relatively higher bargaining power for SaaS customers compared to other software models [19, 60]. All these factors present unique challenges to SaaS vendors who must satisfy their customers' requirements for service to keep churn rates at low levels.

If SaaS is to be accepted and continuously used by its costumers, SaaS vendors need to shift their focus to all relevant aspects of service quality management—that is, all cues and events that occur before, during, and after the delivery of software services. To deliver superior service quality, managers of companies with SaaS offerings must therefore understand how costumers perceive and evaluate SaaS-based services. In this way, they will know in which area to allocate investments to improve their service quality and to increase continued SaaS usage. In order to achieve that goal, it is imperative that they have at their disposal a standardized but complete measurement instrument for assessing SaaS service quality perceptions by their customers.

In this paper, we first explain why existing service quality instruments are insufficient for measuring SaaS service quality. We then describe a study where we develop, refine, test, and validate a service quality measure specifically for SaaS products, which we call "SaaS-Qual." In developing this instrument, our study provides several contributions. First, we answer the call made in a recent article by Bardhan et al. [6] for more experimental and behavioral approaches in service science research, specifically to answer questions on customer experience, such as the following:

> How does the nature of customer contacts change as firms move to more "high-touch" experiential customer service environments? What are the linkages between customer service design and outcomes related to emotional responses, such as customer satisfaction? [6, p. 36]

Second, we continue in the long tradition of service quality research by developing, refining, and testing a service quality instrument (SaaS-Qual) for a specific context (SaaS). As we discuss throughout this paper, although there are already several instruments on service quality in information systems (IS) research, they have mainly focused on service quality assessments in the offline/physical world, such as IS-adapted SERVQUAL [46, 47], or in the business-to-consumer (B2C) e-commerce context, such as E-S-Qual [65] and Web-Qual [54]. Service quality instruments that have addressed software services provided through the Internet, such as ASP-Qual [55, 79], do not fully capture the new issues that arise in a software usage environment that is heavily dependent on an increasingly pervasive cloud infrastructure.

Third, we enhance the existing literature on continued IS usage by providing a detailed conceptualization of the notion of service quality in the context of SaaS. Although previous models on continued IS usage have examined the influence of software service quality confirmation on satisfaction and of continued IS usage intentions [9, 52, 53], they used rather abstract notions of service quality, which is highly desirable for theory-building purposes. Within the context of SaaS and in order to offer more diagnostic and thus prescriptive advice, we provide a more in-depth conceptualization of SaaS service quality, which offers more insights into where the strengths and weaknesses of SaaS services are that may explain dissatisfaction and possible discontinuance of SaaS usage.

The remainder of this paper is organized as follows. In the next section, we develop the theoretical basis for this work, drawing on prior service quality research in the IS, e-commerce, and marketing literature, and we explain why a service quality measurement instrument that is specific to SaaS is needed. We then discuss the development and construct validity testing of the SaaS-Qual instrument. We end with a discussion of results, highlighting their implications for both research and practice and pointing out limitations and promising areas for future research.

Theoretical Background

Service Quality Research

Extensive research on traditional (non-Internet) service quality has been conducted during the past 25 years. Early scholarly writings suggested that service quality stems from a comparison of what customers feel a company should offer (i.e., their expectations) with the company's actual service performance [51, 62]. Using insights from these studies as a starting point, Parasuraman et al. [61] 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 service quality. He defined service quality as the conformance to customer requirements in the delivery of a service [63]. It is a perceived judgment that results from comparing customer expectations with the level of service that customers perceive to have received [62].

The original SERVQUAL instrument [62, 63] identified five service quality dimensions:

1. Tangibles: This dimension deals with the physical environment. It relates to customer assessments of the facilities, equipment, and appearance of those providing the service.
2. Reliability: This dimension deals with customer perceptions that the service provider is providing the promised service in a reliable and dependable manner, and is doing so on time.
3. Responsiveness: This dimension deals with customer perceptions about the willingness of the service provider to help the customers and not shrug off their requests for assistance.
4. Assurance: This dimension deals with customer perceptions that the service provider's behavior instills confidence in them through the provider's courtesy and ability.

5. Empathy: This dimension deals with customer perceptions that the service provider is giving them individualized attention and has their best interests at heart.

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 [34, 42, 46, 90]. Researchers and practitioners have emphasized SERVQUAL's diagnostic power and thus its practical relevance for management decisions [42, 46, 70, 90]. However, a source of concern has been 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. As a response to this, Kettinger and Lee [46, 47] tested and validated an alternative instrument adapted from marketing referred to as the "zones of tolerance" (ZOT) service quality measure [64]. This new instrument recommended using two different comparison norms for service quality assessment: desired service (the level of service a customer believes can and should be delivered) and adequate (minimum) service (the level of service the customer considers acceptable). Separating these two levels is a ZOT 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 (i.e., between desired and adequate service), rather than a single point (i.e., expected service). In this conceptualization, meeting or missing the ZOT thus represents a positive or negative service quality assessment.

Through the emergence of the Internet and electronic channels, several adaptations to the IS SERVQUAL measurement instrument have been proposed in order to address the apparent differences between the evaluative processes for offline and online service quality. In general, researchers found that studying online service quality requires scale development that extends beyond merely adapting offline scales. Gefen [33], for example, extended the SERVQUAL conceptualization to the electronic context and found that the five service quality dimensions collapse to three for online service quality: (1) tangibles; (2) a combined dimension of responsiveness, reliability, and assurance; and (3) empathy. On the basis of a comprehensive review and synthesis of the extant literature on online service quality, Parasuraman et al. [65] detailed four broad sets of criteria as relevant to electronic service quality perceptions: (1) efficiency, (2) system availability, (3) fulfillment, and (4) privacy. Later, Swaid and Wigand [85] developed a scale for measuring service quality for online retailers that resulted in six dimensions, only some of which overlapped with previous online service quality measures.

In addition, drawing on previous research in IS SERVQUAL and the online service quality literature, a few IS researchers have transferred and adapted the findings to the application service provider (ASP) context where software applications are not installed on the client's in-house servers, but delivered over a network. In an exploratory study, Ma et al. [55] 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. Also, Sigala [79] developed an ASP service quality model for companies evaluating their ASP-hosted online Web stores. Her analysis suggested a multifactor scale, including tangibles, reliability, responsiveness, assurance, empathy, trust, business understanding, benefit and risk share, conflict, and commitment.

Although several facets from these previous studies can be transferred and adapted to the SaaS context, some factors that are important for SaaS service quality are not addressed. For example, issues of flexibility and security, which are at the core of SaaS services [56, 60, 73], are insufficiently covered in previous studies. In addition, the ASP service quality scales developed by Ma et al. [55] and Sigala [79], which are closest to SaaS from a conceptual point of view, have not been adequately validated with a confirmatory factor analysis and a nomological validity test. Table 1 presents a detailed comparison of all the existing studies of adaptations of IS SERVQUAL, including more details on the extent to which each prior measure would be suitable or not for measuring SaaS service quality.

A Comparison of SaaS and ASP

Although the SaaS and ASP models of software delivery share some similarities, they also have significant differences that result in the need for a service quality measure that is specific to SaaS (refer to Table 2 for an in-depth analysis of the differences between SaaS and ASP). In general, contrary to the ASP single-tenant architecture, SaaS uses a multi-tenant architecture, where 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 [2, 94]. Customer-specific configuration can be made only at the meta-data layer on top of the common code using an interface provided by the SaaS vendor. The service can be integrated with other applications, or it can connect with more custom functions through common Web services application programming interfaces (APIs) that are defined and maintained by SaaS vendors [94]. This new architecture has important implications for SaaS customers' service quality perceptions, which can differ from those of ASP customers, therefore rendering ASP-specific measures of service quality insufficient:

- First, software applications and IT infrastructure (i.e., servers, storage, and bandwidth) are shared across customers in the SaaS model, contrary to the classical ASP model, where they are dedicated to each single customer. As a result, there are different client expectations regarding system performance, availability, and security aspects. For example, higher network bandwidth and processing power in the SaaS model has increased customer expectations of reliability and responsiveness as compared to the ASP model.
- Second, the SaaS model constrains client options for customization of the main functionality and data structures of the software. In the ASP model, on the other hand, due to its single-tenant nature, customers can have higher expectations regarding customized services.
- Third, the SaaS model gives more control over future development to the vendor, as clients have no choice but to adopt future upgrades of software if they continue

Table 1. Prior SERVQUAL Studies and Their Applicability in the SaaS Context

Source	SERVQUAL instrument	Factors	Research context	Validation of scale	Applicability to measure SaaS service quality
Parasuraman et al. [63]	SERVQUAL	Tangibles, reliability, responsiveness, assurance, empathy	Offline, B2C (service and retailing organizations)	EFA, CFA, SA	Applied only in the offline context missing facets important for online environments (e.g., security regarding data transmission, availability of an Internet connection); not adapted to IS.
Kettinger and Lee [45, 46, 47] based on Zeithaml et al. [95]	(Zone-of-tolerance-based) IS-SERVQUAL	Reliability, responsiveness, rapport (assurance/empathy), tangibles	Offline, IS function in companies/universities, offline	EFA, CFA	Although the items of SERVQUAL were adapted to the IS realm, they were only applied in the offline context missing facets important for online environments (e.g., security regarding data transmission, availability of an Internet connection).
Jiang et al. [42], Pitt et al. [69], Watson et al. [90]	IS service quality	Tangibles, reliability, responsiveness, assurance, empathy	Offline, IS department in companies (e.g., financial institution)	EFA, CFA	Similar to Kettinger and Lee [47], these scales were applied only in the offline context to measure service quality of IS departments missing factors/facets important for measuring service quality of external service providers in online environments.
Gefen [33]	SERVQUAL in e-commerce	Tangibles, empathy, reliability/responsiveness/assurance	E-commerce Web sites, B2C	EFA, CFA, SA	Adapted previous scales [62, 90] to a B2C e-commerce context; although several items can be transferred to the SaaS context (which is also Internet based), important facets are missing to fully capture SaaS service quality (e.g., security or flexibility facets).

(continues)

Table 1. Continued

Source	SERVQUAL instrument	Factors	Research context	Validation of scale	Applicability to measure SaaS service quality
Wolfinbarger and Gilly [93]	eTailQ	Fulfillment/reliability, Web site design, privacy/security, customer service	E-commerce Web sites, B2C	EFA, CFA, SA	Besides items referring to traditional factors (e.g., reliability, responsiveness), the scale also introduces some security facets that can be transferred to the SaaS context; however, the scale items are tailored to a B2C online-shopping context with physical products being sold (purchase transactions) and does not capture SaaS-specific items (e.g., specific SaaS features such as reporting or dashboard features).
Parasuraman et al. [65]	E-S-Qual (electronic service quality)	Efficiency, fulfillment, system availability, privacy	E-commerce Web sites, B2C	EFA, CFA, SA	The scale items were specifically developed for a B2C e-commerce setting and miss important aspects of SaaS service quality (e.g., data/network security or flexibility); however, some aspects and items are also transferable to the SaaS context (e.g., system availability as part of responsiveness).
Barnes and Vidgen [7], Loiacono et al. [54]	WebQual	Informational fit-to-task, tailored information, trust, response time, ease of understanding, intuitive operations, visual appeal, innovativeness, emotional appeal, consistent image, online completeness, relative advantage	(Commercial, organizational) Web sites, B2C	EFA, CFA, SA	WebQual is a measure to evaluate organizational Web sites from a consumer perspective; while tapping into important facets on usefulness and ease of understanding of a Web site, important facets are missing to capture facets that are more specific to the SaaS context (e.g., security or flexibility facets).

Swaid and Wigand [85]	Quality of E-service	Web site usability, information quality, reliability, responsiveness, assurance, personalization	E-commerce Web sites, B2C	EFA, CFA, SA	Measures the quality of e-service experience including cues that occur before, during, and after the e-purchase transaction; although some aspects are also relevant for the SaaS context (e.g., responsiveness or reliability items), important facets of SaaS service quality are not covered by this scale (e.g., security or flexibility facets).
Sigala [79]	ASP-Qual	Tangibles, reliability, responsiveness, assurance, empathy, trust, business understanding, benefit and risk share, conflict, commitment	ASP, B2B	EFA only	The scale has not been tested empirically with CFA, and SA and does not capture customer expectations and performance perceptions in a ZOT-based format; although many relevant facets of ASP-Qual also apply to SaaS, some important facets are missing, such as security and flexibility that are highly relevant for assessing SaaS service quality.
Ma et al. [55]	Factors of service quality for ASP	Features, availability, reliability, assurance, empathy, conformance, security	ASP, B2B	EFA only	The scale has not been tested empirically based on CFA and SA; although security has been added to this scale, the suggested items do not cover all relevant security aspects that are relevant for SaaS; flexibility is also missing in the final scale as a crucial facet for evaluating SaaS service quality.

Notes: EFA = exploratory factor analysis, CFA = confirmatory factor analysis, SA = structural analysis (e.g., structural equation modeling).

Table 2. The Differences Between ASP and SaaS and Their Implications on Service Quality

Aspects/sources for potential differences	ASP	SaaS	Potential changes in service quality expectations regarding SaaS [2, 8, 18, 56, 57, 60, 73, 94]
Network bandwidth and processing power	Low (speed of applications was oftentimes unacceptable)	High (bandwidth and processing power are no longer constraints, allowing for higher speed and less errors)	Customer expectations regarding *reliability* and *responsiveness* of software services have increased due to the availability of higher network bandwidth and processing power.
IT infrastructure and architecture	Single-tenancy (IT infrastructure and application are dedicated to individual clients)	Multi-tenancy (based on virtualization and load balancing technologies, IT infrastructure and application are shared across clients)	Customer expectations regarding *rapport* (assurance/ empathy) may have increased because service providers have to balance the needs/requirements across different customers that now share the same infrastructure and application code. Scalability/elasticity of resources is easier to realize in SaaS, as multi-tenancy is based on load balancing/ virtualization, i.e., users/licenses can more easily be added or deleted, providing higher *flexibility.*
Level of standardization of application software across customers	Middle (applications usually used in an ASP contract could be customized by clients. However, the applications had rather monolithic architectures, often based on their on-premise counterpart)	High (applications are increasingly based on platform ecosystems with highly standardized modules and clearly defined interfaces to flexibly add complementary functionalities from third-party developers)	Customization of *existing* software modules to individual customer needs is harder to realize in SaaS, because customers share a common, standardized code base. As SaaS operations are more standardized across customers, individual attention to customer needs may suffer leading to lower *rapport* (empathy/assurance).

Provision of third-party software modules (extensibility)	Low (due to the monolithic nature of ASP software, the provision of software modules from third parties to extend the core functionality is very restricted)	High (SaaS is based on a cloud infrastructure that comprises standardized interfaces and a common code base, also called platform, that is open to third-party developers)	Since SaaS applications are most often based on an open platform over which third-party developers can provide additional functionality (e.g., Salesforce .com's platform called Force.com), customers expectations regarding *flexibility* of extending the core functionality of the SaaS application may have changed, so that more modules are available that can be added or combined with existing ones.
Importance of security for customers	Middle (individual, dedicated solutions are easier to protect)	High (shared SaaS/cloud services are more open and easier to attack)	Customer expectations regarding *security* may have increased in the SaaS realm due to its more open nature.
Ease of use for application users (Integration into Web 2.0 usage environment)	Low (ASP applications were usually based on user interfaces from their on-premise counterparts)	High (SaaS applications resemble Internet user interfaces that users are familiar with and thus can transfer their experiences)	Customers may expect user interfaces and Web site *features* that resemble typical Internet features that are integrated with other interactive features such as search or social media features.
Interoperability with existing customer systems	Low (very difficult due to lack of standardized interfaces)	High (more standardized interfaces are available with a minimum of changes to customer systems)	Due to easier integration with customer systems, customer expectations regarding *flexibility* of connecting different applications, *reliability* of the entire application architecture, and *features* drawing on functionality from different applications may have changed.

using the service. This may result in increased demands on the vendors for reliable and frequent upgrades, as opposed to the ASP model where the customers may bear some of that responsibility themselves. In a similar vein, 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 customized interface [94]. Thus, this model no longer requires any client-specific investment by the vendor and helps vendors to reap significant economies of scale. They can share one application cost effectively across hundreds of companies, which is a vast improvement on the old ASP model [60].

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 Web browsers [89]. But SaaS skeptics point out that the limited customization possibilities of SaaS applications and potential traffic bottlenecks due to sharing IT infrastructure across many customers may hamper service quality dimensions [8]. Given that SaaS presents unique service quality challenges for vendors and therefore unique service quality expectations for their clients, we believe it is necessary to develop a measurement instrument to capture the specific facets of service quality in SaaS.

Developing the SaaS-Qual Instrument

In adherence to established scale development guidelines [20, 27, 37] and key components of validity (i.e., content, face, discriminant, convergent, nomological, and prescriptive validity) [5, 82], we used a systematic three-step process, involving a variety of methods to develop, refine, and validate the measurement of SaaS service quality (SaaS-Qual). As shown in Figure 1, the three steps were (1) conceptual development and initial item pool generation, (2) conceptual refinement, item modification, and pilot study, (3) main study and validity testing. Table 3 shows the major roles and results of research participants in developing the SaaS-Qual measure throughout the three stages.

Step 1: Conceptual Development and Initial Item Pool Generation

The first step of the process involved the development of a beginning set of factors for SaaS-Qual along with a pool of measurement items for those factors. This was accomplished with a two-pronged approach. First, we performed a comprehensive review of prior studies in the IS, e-commerce, and marketing literature. Practitioner-oriented and popular press publications were also examined to ensure that no factor was overlooked because of the newness of the SaaS model. A number of factors and measurement items from prior studies were included without changes, whereas others were adapted from

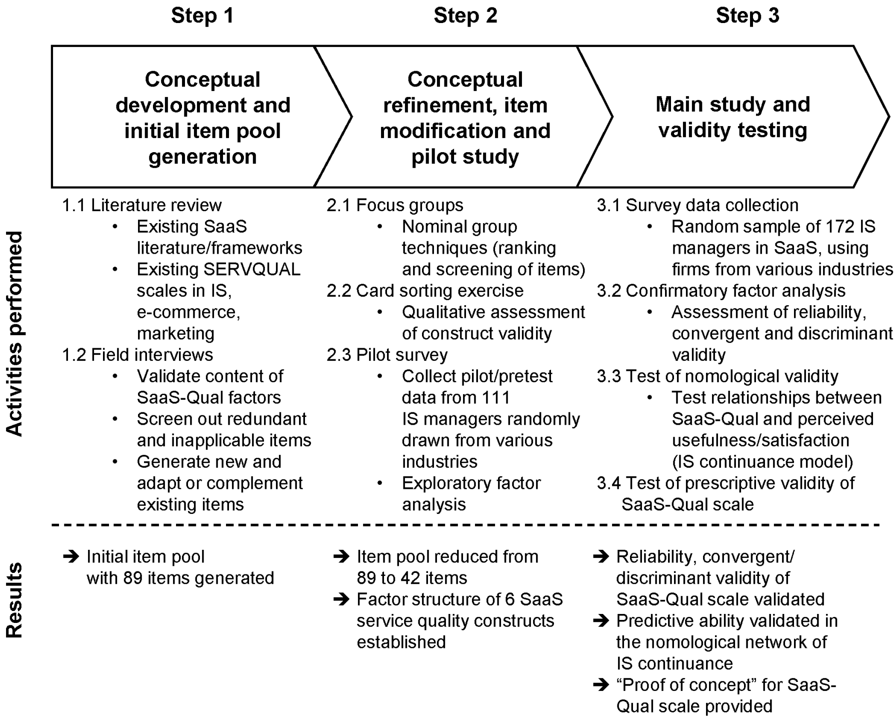

Figure 1. Overview of the Scale Refinement Process

previous service quality scales (i.e., ZOT-based IS-SERVQUAL [47], E-S-Qual [65], Quality of E-Service [85], and ASP-Qual [55, 79]).

The second part of our approach involved in-depth field interviews with relevant experts. Ten interviews were conducted with account managers of SaaS providers who had, on average, more than 12 years of work experience and had experience with SaaS-based applications, on average, for more than 5 years. Since account managers of SaaS providers are concerned with what their SaaS customers want and have developed criteria based on these desires, this group of experts seemed appropriate for capturing all relevant content aspects of SaaS service quality. The specific goal of the field interviews was to validate the beginning set of the SaaS-Qual factors, to screen out redundant or inadequate items and produce new ones in order to ensure completeness (i.e., that no key factors were overlooked) and content validity of the scales.

As a result of the field interviews, the factors *Assurance* and *Empathy* merged into one factor that we called *Rapport,* similar to Kettinger and Lee [47]. *Reliability* and *Responsiveness* were validated as two separate factors. Similar to Ma et al. [55], *Tangibles* was renamed *Features.* In other words, the importance of the four established service quality factors (i.e., *Rapport, Responsiveness, Features,* and *Reliability*) in the SaaS context was confirmed. In addition, the IS managers we interviewed consistently emphasized the relevance of two additional distinct factors for a complete SaaS-Qual measurement instrument, namely *Flexibility* and *Security,* which we added to the other

Table 3. Roles of Research Participants

Steps	Participants	Number of participants	Activities/roles	Results
Field interviews	Account managers of SaaS providers	10	Initial validation of the SaaS-Qual factors, screening out of redundant and inapplicable items and generation of additional items	Conceptual framework for SaaS-Qual factors, initial item pool with 89 items
Focus group discussion	IS managers (SaaS using companies), IS researchers	25 (14, 11)	Generation of a ranked list of items; screening out redundant, inapplicable, or low-ranked items; rewording items to improve clarity (face validity); feedback on initial questionnaire draft	Initial item pool reduced to 63 items
Card-sorting procedure	IS managers (SaaS using companies), IS researchers	6 (3, 3)	Assessment of content validity of items using a card-sorting exercise, screening out ambiguous items	Reduction of items from 63 to 50 items
Pilot study	IS managers (SaaS using companies)	111	Reduction of items and refinement of survey instrument	111 valid responses, 50 items reduced to 42, editorial changes to survey
Main study	IS managers (SaaS using companies)	172	Response to the final survey instrument	172 valid responses

Table 4. Conceptual Definitions of the Six SaaS-Qual Factors

Factor	Conceptual definition
Rapport	Includes all aspects of an 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., support tailored to individual needs).
Responsiveness	Consists of all aspects of an 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., 24-7 hotline support availability) is guaranteed.
Reliability	Comprises all features of an SaaS vendor's ability to perform the promised services timely, dependably, and accurately (e.g., providing services at the promised time, provision of error-free services).
Flexibility	Covers the degrees of freedom customers have to change contractual (e.g., cancellation period, payment model) or functional/technical (e.g., scalability, interoperability, or modularity of the application) aspects in the relationship with an SaaS vendor.
Features	Refers to the degree the key functionalities (e.g., data extraction, reporting, or configuration features) and design features (e.g., user interface) of an SaaS application meet the business requirements of a customer.
Security	Includes all aspects to ensure that regular (preventive) measures (e.g., regular security audits, usage of encryption, or antivirus technology) are taken to avoid unintentional data breaches or corruptions (e.g., through loss, theft, or intrusions).

four factors. Furthermore, we dropped several items due to redundancy (e.g., *Responsiveness* or *Reliability* items of different previous scales were consolidated) and due to the inappropriateness of some items for the SaaS context (e.g., items referring to the delivery of physical products which is important in the e-commerce context but not in SaaS). Based on the feedback of the interviewees, new items were also developed (e.g., new *Security, Flexibility,* and *Features* items were generated to fully cover the content of SaaS service quality) based on guidelines from scale development literature [37]. This process led to our initial set of six factors (see the conceptual definitions in Table 4) with an initial measurement item pool of 89 items (our initial list of 89 items is omitted here for brevity but can be obtained from the authors).

Step 2: Conceptual Refinement, Item Modification, and Pilot Study

The six SaaS-Qual factors and the initial pool of 89 measures that resulted from Step 1 were then refined and modified through a focus group discussion, a card-sorting procedure, and a pilot study.

Focus Group Discussion

For the focus group discussion, we followed an approach similar to the one used by Davis [24]. Twenty-five participants were invited, including 14 IS managers of SaaS-using companies having, on average, more than 4 years of experience with SaaS implementations, and 11 IS researchers with special expertise on IT services management and research. The specific goals of the focus group discussion were to (1) let participants independently rank the 89 items on how close they were in meaning with that of the underlying SaaS-Qual factors using nominal group techniques [80]; (2) eliminate redundant, inapplicable, or low-ranked items; (3) reword items to improve clarity (i.e., face validity); and (4) obtain feedback on the length, format, and clarity of the instructions and initial questionnaire draft. On the basis of insights from the focus group discussion, we reduced the initial set of 89 items to 63 by screening out redundant, inapplicable, and low-ranked items and reworded them to improve their face validity. We then included the 63 items in a card-sorting exercise.

Card-Sorting Exercise

To support construct validity and to evaluate how well the 63 items tapped the six SaaS-Qual factors, we conducted a card-sorting exercise with the assistance of six judges [37, 59]. These judges included 3 IS managers and 3 IS researchers (4 male, average age = 43, average of 4.5 years of experience with SaaS). Each judge was presented with the 6 SaaS-Qual factors, the definitions for each factor, and a randomly sorted list of the 63 items printed on 3- × 5-inch index cards. The judges were instructed to individually assign each item to 1 of the 6 factors or to an "ambiguous" category if they were unsure of the best placement. After completing the sorting procedure, they explained why they sorted cards (if any) into the "ambiguous" category. We observed that all six judges had difficulties with several specific items either because they were ambiguous or confusing. For example, the judges felt that the item "The company providing SaaS services is reputable" was ambiguous because it tapped several other factors of the SaaS-Qual framework, or the item "The SaaS service provider helps customers set proper expectations" seemed to be too broad to clearly fit into any of the six SaaS-Qual factors. We thus eliminated 13 items that at least 4 out of the 6 judges marked as ambiguous or confusing, ending with 50 items. The card-sorting exercise resulted in an average "hit ratio" of 85 percent across the 6 SaaS-Qual factors, and an average kappa [21] of 0.83, both good indications of construct validity.[1] The 50 items are shown in Table 5.

Pilot Study

To further validate the relevance, coverage, and clarity of the measurement items, we conducted a pilot survey study. We designed a questionnaire including the 50 items that resulted from the card-sorting exercise, which we sent to a random sample of 1,000 German-based companies. The companies were drawn from the

Table 5. Items for the SaaS-Qual Scale, Sources, and Refinement Results (Pilot Study)

SaaS-QUAL factors (abbreviation)	Sources	Items*	Refinement result
		When it comes to . . .	
Rapport (Ra)	Ra1 [55]	. . . a shared approach to problem solving	Final item
	Ra2 [55]	. . . customer-specific trainings and courses	Final item
	Ra3 [79]	. . . understanding our business goals and processes	Final item
	Ra4 (new)	. . . a good personal relationship	Final item
	Ra5 (new)	. . . an aligned working style (e.g., convenient operating hours)	Final item
	Ra6 [79]	. . . having the knowledge to answer customers' questions	Final item
	Ra7 [79]	. . . the cultural fit between SaaS provider and our company	Final item
	Ra8 [79]	. . . support that is tailored to our individual needs	Final item
	Ra9 [47]	. . . having the customer's best interest at heart	Final item
	Ra10 [79]	. . . giving users individual attention	Deleted item
	Ra11 [47]	. . . instilling confidence in customers	Deleted item
Responsiveness (Res)	Res1 [65]	. . . system availability/uptime for business (e.g., system crash or freeze)	Final item
	Res2 [55]	. . . network performance	Final item
	Res3 [55]	. . . efficient disaster recovery	Final item
	Res4 [55]	. . . efficient contingency and replacement policy	Final item
	Res5 [55]	. . . hardware and software redundancy	Final item
	Res6 (new)	. . . an adequate number of service personnel dedicated to our company	Final item
	Res7 [47]	. . . the support of up-to-date, cutting-edge hardware, software, and netware technology	Final item
	Res8 [47]	. . . (technical) support availability (i.e., the promptness of providing services)	Final item
	Res9 [47]	. . . (multichannel) customer care (i.e., the SaaS provider's willingness to help users)	Final item
	Res10 [79]	. . . telling users exactly when services will be performed	Deleted item
	Res11 [85]	. . . providing inbound and outbound systems to deal with customer complaints	Deleted item

(continues)

Table 5. Continued

SaaS-QUAL factors (abbreviation)	Sources	Items*	Refinement result
		When it comes to . . .	
Reliability (Rel)	Rel1 [47]	. . . providing services at the promised time	Final item
	Rel2 [47]	. . . performing services right the first time	Final item
	Rel3 (new)	. . . fulfilling the obligations to the contract (including SLAs)	Final item
	Rel4 [79]	. . . user problems, our SaaS provider shows a sincere interest in solving them reliably	Final item
	Rel5 [79]	. . . the provision of error-free services and accurate budgetary controls	Final item
	Rel6 [47]	. . . providing appropriate quality certifications	Deleted item
Flexibility (Fl)	Fl1 [55]	. . . the integration and interoperability of the SaaS application with our information and communications technology infrastructure	Final item
	Fl2 [55]	. . . application scalability (i.e., number of user subscriptions)	Final item
	Fl3 (new)	. . . modularity of features from which can be chosen (i.e., packaging choices)	Final item
	Fl4 (new)	. . . having access to the latest software versions	Final item
	Fl5 [55]	. . . modifying contractual parameters at later stages	Final item
	Fl6 [55]	. . . choices of ways to pay (e.g., payment/billing options)	Final item
	Fl7 (new)	. . . application customization (i.e., configurability)	Deleted item

Features (Fe)	Fe1 [55, 79]	. . . a visually appealing and sympathetic user interface	Final item
	Fe2 [79]	. . . a user-friendly navigation structure and search functionality	Final item
	Fe3 [55]	. . . data reporting and extracting features	Final item
	Fe4 [55]	. . . the SaaS application's configuration (e.g., user administration, etc.) features	Final item
	Fe5 (new)	. . . the SaaS application's help functionalities	Final item
	Fe6 (new)	. . . the dashboard features with metrics measuring customers service usage	Final item
	Fe7 (new)	. . . the SaaS application's core features to support process steps/activities	Final item
	Fe8 [55]	. . . application performance monitoring	Deleted item
Security (Sec)	Sec1 (new)	. . . data backup and recovery	Final item
	Sec2 [55]	. . . regular security audits	Final item
	Sec3 [55]	. . . providing a secure physical environment (i.e., secure data center)	Final item
	Sec4 [55]	. . . anti-virus protection	Final item
	Sec5 [55]	. . . data encryption	Final item
	Sec6 [55]	. . . data confidentiality	Final item
	Sec7 (new)	. . . access control measurements	Deleted item

* Rated on 7-point Likert scales with 1 = "low" and 7 = "high" for "My minimum service level is . . . ," "My desired level of service is . . . ," and "My perceived service level is . . ." [47])

Hoppenstedt firm database (Bisnode Business Information Group), which is one of the largest commercial business data providers in Germany. The database contains over 300,000 profiles of German companies, their branches, and the major industrial associations in Germany. To support the external validity of our study, we did not constrain the sample to specific industries or to firms of a specific organizational size. The survey questionnaire was mailed to the highest-ranking IS manager in each 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 or her firm position compares to firms of similar size and in similar industries.

After 41 responses were dropped due to missing data, a total of 111 usable responses were deemed usable for our analysis. To assess potential threats of nonresponse bias for the sample, 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.

As a preliminary check of data quality, skewness and kurtosis scores for each item were analyzed. All of these scores were within the –2 to +2 range (skewness: –1.821 to 1.363; kurtosis: –0.841 to 1.434), suggesting no serious deviations from the normality assumption [35]. Furthermore, we screened the data set for outliers using Cook's D and standardized residuals but did not detect any outlier cases [10]. Using the data from this initial pilot study, we conducted an exploratory factor analysis with SPSS 17.0, using principal components analysis with oblique (i.e., direct oblimin) rotation. We then went through a series of iterations, each involving elimination of items with low loadings on all factors or high cross-loadings on two or more factors, followed by factor analysis of the remaining items. This iterative process resulted in the final SaaS-Qual instrument, consisting of 42 items that tapped into the 6 SaaS-Qual factors (see Table 6). Using this final 42-item instrument, we tested the unidimensionality/convergent validity as well as the composite reliability of the six SaaS-Qual factors using LISREL version 8.8 [11, 43].

Unidimensionality and convergent validity require that one single latent variable underlies a set of measures [1]. To test unidimensionality and convergent validity, we generated six first-order models with each corresponding to one of the six factors of SaaS-Qual. The results shown in Table 7 suggest that all six SaaS-Qual factors demonstrated adequate levels of model fit. Overall, the results indicate that the measures of each of the six SaaS-Qual factors satisfy the unidimensionality and convergent validity requirements. The composite reliability (p_c), which represents the proportion of measure variance attributable to the underlying latent variable, was calculated to assess the reliability of the measure [92]. Values of p_c in excess of 0.50 indicate that the variance captured by the measures is greater than that captured by error components [3]. All of the six SaaS-Qual factors had p_c values between 0.87 and 0.93, thus suggesting satisfactory levels of reliability.

Table 6. EFA Results for the SaaS-Qual Scale (Pilot Study)

	Range of EFA loadings (pattern matrix)[a]					
Factors[b]	Rapport	Responsiveness	Reliability	Flexibility	Features	Security
Rapport						
Ra1	**0.93**	0.25	0.15	0.21	0.26	0.19
Ra2	**0.91**	0.15	0.18	0.20	0.27	0.12
Ra3	**0.94**	0.26	0.25	0.25	0.23	0.22
Ra4	**0.92**	0.24	0.21	0.17	0.21	0.15
Ra5	**0.91**	0.20	0.22	0.25	0.30	0.15
Ra6	**0.90**	0.21	0.24	0.21	0.34	0.09
Ra7	**0.91**	0.23	0.23	0.17	0.24	0.17
Ra8	**0.93**	0.26	0.24	0.23	0.30	0.17
Ra9	**0.89**	0.18	0.20	0.29	0.34	0.23
Ra10	~~**0.51**~~	~~0.32~~	~~0.40~~	~~0.23~~	~~0.09~~	~~0.14~~
Ra11	~~**0.38**~~	~~0.28~~	~~0.35~~	~~0.19~~	~~0.13~~	~~0.19~~
Responsiveness						
Res1	0.28	**0.88**	0.15	0.32	0.12	0.32
Res2	0.25	**0.89**	0.22	0.31	0.12	0.22
Res3	0.25	**0.89**	0.19	0.29	0.19	0.23
Res4	0.25	**0.91**	0.26	0.28	0.15	0.18
Res5	0.17	**0.93**	0.22	0.26	0.16	0.16
Res6	0.19	**0.92**	0.27	0.27	0.09	0.18
Res7	0.25	**0.93**	0.21	0.28	0.12	0.17
Res8	0.23	**0.90**	0.17	0.27	0.14	0.19
Res9	0.21	**0.89**	0.18	0.30	0.09	0.23
Res10	~~0.35~~	~~**0.45**~~	~~0.24~~	~~0.41~~	~~0.29~~	~~0.41~~
Res11	~~0.23~~	~~**0.41**~~	~~0.32~~	~~0.34~~	~~0.29~~	~~0.40~~

(continues)

Table 6. Continued

	Range of EFA loadings (pattern matrix)[a]					
Factors[b]	Rapport	Responsiveness	Reliability	Flexibility	Features	Security
Reliability						
Rel1	0.29	0.27	**0.90**	0.15	0.30	0.12
Rel2	0.30	0.28	**0.89**	0.19	0.30	0.11
Rel3	0.32	0.27	**0.87**	0.21	0.28	0.17
Rel4	0.32	0.29	**0.84**	0.22	0.25	0.23
Rel5	0.28	0.24	**0.80**	0.20	0.23	0.20
Rel6	~~0.31~~	~~0.41~~	~~**0.47**~~	~~0.09~~	~~0.46~~	~~0.14~~
Flexibility						
Fl1	0.27	0.42	0.15	**0.83**	0.13	0.23
Fl2	0.28	0.44	0.17	**0.82**	0.22	0.17
Fl3	0.20	0.36	0.22	**0.79**	0.21	0.05
Fl4	0.23	0.39	0.08	**0.78**	0.11	0.25
Fl5	0.27	0.41	0.18	**0.84**	0.13	0.18
Fl6	0.29	0.40	0.27	**0.74**	0.24	0.08
Fl7	~~0.21~~	~~0.39~~	~~0.28~~	~~**0.47**~~	~~0.44~~	~~0.23~~

Features						
Fe1	0.41	0.13	0.24	0.19	**0.88**	0.12
Fe2	0.35	0.13	0.27	0.15	**0.87**	0.07
Fe3	0.39	0.13	0.26	0.18	**0.89**	0.09
Fe4	0.39	0.15	0.22	0.16	**0.87**	0.13
Fe5	0.38	0.13	0.24	0.12	**0.88**	0.09
Fe6	0.37	0.09	0.26	0.16	**0.92**	0.14
Fe7	0.34	0.10	0.25	0.15	**0.89**	0.13
Fe8	~~0.46~~	~~0.08~~	~~0.21~~	~~0.09~~	**~~0.49~~**	~~0.23~~
Security						
Sec1	0.28	0.31	0.19	0.15	0.10	**0.82**
Sec2	0.21	0.34	0.18	0.19	0.12	**0.80**
Sec3	0.19	0.30	0.17	0.16	0.13	**0.79**
Sec4	0.21	0.29	0.20	0.24	0.16	**0.78**
Sec5	0.22	0.28	0.19	0.23	0.09	**0.81**
Sec6	0.24	0.32	0.16	0.24	0.12	**0.82**
Sec7	~~0.19~~	~~0.34~~	~~0.38~~	~~0.12~~	~~0.41~~	**~~0.46~~**

Notes: EFA = exploratory factor analysis. Boldface values indicate primary (target) factor loadings, non-boldface values indicate item cross-loadings. [a] Total variance extracted by the six factors = 89 percent; rotation method: direct oblimin with Kaiser normalization; average interfactor correlation = 0.45; selection criteria: factor loading ≥ 0.5, no multiple loadings, items not meeting these criteria were dropped as indicated by the struck-out lines. [b] All items are measured as difference score between perceived and minimum accepted service levels [47].

Table 7. Convergent Validity/Composite Reliability Results (Pilot Study)

Factor	Number of items	χ^2	df	χ^2/df (< 5)	GFI (> 0.90)	AGFI (> 0.80)	RMR (< 0.10)	CFI (> 0.90)	p_c
Rapport	9	98.25*	27	3.64	0.92	0.84	0.09	0.91	0.88
Responsiveness	9	114.68*	27	4.25	0.90	0.82	0.10	0.90	0.93
Reliability	5	12.51*	5	2.50	0.96	0.87	0.07	0.98	0.87
Flexibility	6	26.57*	9	2.95	0.91	0.81	0.08	0.93	0.91
Features	7	58.23*	14	4.16	0.90	0.81	0.10	0.96	0.92
Security	6	12.35*	9	1.37	0.96	0.91	0.06	0.99	0.91

Notes: df = degrees of freedom; GFI = goodness-of-fit index; AGFI = adjusted goodness-of-fit index; RMR = root mean square residual; CFI = comparative fit index.
* $p < 0.05$.

Step 3: Main Study and Validity Testing

In order to gather sufficient evidence of convergent, discriminant validity and predictive ability (nomological validity) of the proposed scale, we conducted a main study, using a confirmatory survey of IS managers in SaaS-using firms.

Survey Data Collection

For the purposes of testing the validity of the 42-item SaaS-Qual, including its nomological validity, we developed a new questionnaire that included the 42 items as well as scales measuring the other constructs (i.e., satisfaction, perceived usefulness, and SaaS continuance intention; see Table 8) included in the nomological network, which will be discussed later in the paper. We sent this questionnaire to a random sample of 2,000 companies, different from the sample of 1,000 companies used in the scale refinement step but drawn from the same Hoppenstedt firm database. After sorting out 73 responses due to missing data, we received 172 usable responses, resulting in a response rate of 8.6 percent. As in the first survey, we introduced this second questionnaire 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 most senior IS manager having a good overview of the organization's perception of SaaS service quality. Moreover, to increase the content validity of the responses and avoid social desirability bias, we also asked the respondents to fill out the questionnaire regarding one specific SaaS application type (e.g., enterprise resource planning or customer relationship management) that they were using or were familiar with, and not a typical or very successful one [66, 77]. Similar to the first survey, we compared the respondent and nonrespondent firms with respect to sales, industry, and the number of employees. We found no significant differences ($p > 0.05$).

More than 40 percent of the firms in the research sample had had SaaS in use for more than 3 years. Almost half the firms (46.5 percent) had 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 percent), wholesale/retail trade (25.00 percent), financial intermediation (15.12 percent), TIME (telecommunication, information technology, media, entertainment) industries (11.05 percent), construction and real estate (8.72 percent), logistics (5.23 percent), public health care (4.07 percent), and electricity/gas/water supply (1.74 percent).

The main study (as well as the pilot study) utilized a "key informants" methodology for data collection, where 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 [78]. 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 not only to identify organizations that actively used SaaS applications but also to identify respondents within those organizations who were intimately involved with, and most knowledgeable about,

Table 8. Items for Other Scales in Main Study

Constructs	Indicators	
SaaS continuance intention (SCI) [9]	SCI1	We intend to continue using the SaaS-based software delivery model rather than discontinue its use.
	SCI2	Our intentions are to continue using the SaaS-based software delivery model than use any alternative means (e.g., on-premise solutions).
	SCI3	If I could, I would like to discontinue the use of the SaaS-based software delivery model (reverse coded).
Satisfaction (S) [9]	How do you feel about your overall experience of using the SaaS-based software delivery model?	
	S1	Very dissatisfied/very satisfied.
	S2	Very displeased/very pleased.
	S3	Very frustrated/very contented.
	S4	Absolutely terrible/absolutely delighted.
Perceived usefulness (PU) [25]	PU1	Using the SaaS-based software delivery model (compared to other ones) improves our performance in managing our functions/processes.
	PU2	Using the SaaS-based software delivery model (compared to other ones) increases our productivity in managing our functions/processes.
	PU3	Using the SaaS-based software delivery model (compared to other ones) enhances our effectiveness in managing our functions/processes.
	PU4	Overall, the SaaS-based software delivery model (compared to other ones) is useful in supporting our functions/processes.

SaaS usage. Demographic information about the respondents showed that about 47.1 percent were senior IT executives and 39.5 percent were information technology (IT) managers. Although some preliminary steps were taken to ensure appropriate selection of key informants, a formal check was administered as part of the questionnaire [50]. Specifically, two items were used to assess an informant's length and frequency of SaaS usage. The mean score for the length of SaaS usage was 24.42 months and for the frequency of SaaS usage 20.56 times a month, indicating that respondents were appropriate. Respondents from larger companies and from companies with a higher number of years since organizational SaaS rollout had used SaaS significantly longer and more often in the past ($p < 0.05$). Further, IT (middle) managers had used SaaS significantly longer and more often than IT and business executives ($p < 0.05$). Further sample characteristics are shown in Table 9.

Because the data were obtained from one key respondent from every organization, we conducted Harman's one-factor test to ensure there was no significant presence of common method bias [71]. 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 (i.e., the most covariance explained by one factor is 16.03 percent), indicating that common method biases are not a likely contaminant

Table 9. Sample Descriptives for Main Study ($N = 172$)

Category	Percent
Number of employees	
< 49	30.2
50–500	39.5
> 500	30.2
Annual sales (millions of euros)	
< 5	28.5
5–50	34.3
> 50	37.2
Number of years since SaaS rollout	
< 1	13.4
1–3	46.5
3–5	32.6
> 5	7.6
Respondent title	
IT executives (chief information officer/chief technology officer/vice president of IS/IT)	47.1
Business executives (chief executive officer, chief financial officer, and chief operating officer)	9.9
IT (middle) managers	39.5
Business managers and users	3.5
Length of SaaS usage of respondent	
< 3 months	6.3
3 to less than 6 months	7.7
6 to less than 12 months	17.6
12 months or more	68.4
Frequency of SaaS usage of respondent	
4 or less times a month	5.7
5 to 8 times a month	15.3
9 to 12 times a month	33.9
13 or more times a month	45.1

of our results. Furthermore, a correlational marker technique was used, in which the highest variable from the factor analysis was entered as an additional independent variable [74]. 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 the SaaS-Qual instrument and its validity in a nomological network.

Convergent and Discriminant Validity Tests

Convergent and discriminant validity and scale unidimensionality of the scales used in the main study were assessed via confirmatory analysis using SmartPLS 2.0 [75]. The choice was motivated by several considerations. Partial least squares (PLS) can be used to estimate models that use both reflective and formative indicators, allows for modeling latent constructs under conditions of nonnormality (which was not the case in the main study, as all skewness and kurtosis estimates ranged within the –2 to +2 range and no outliers were detected), and is appropriate for complex models that include many variables/indicators and relationships despite low to medium sample size [16].[2] We assessed the psychometric properties of the SaaS-Qual measurement models, presented in Table 10, by examining individual item loadings, internal consistency, convergent validity, and discriminant validity.

Convergent validity was evaluated for the six constructs using three criteria recommended by Fornell and Larcker [32]: (1) all measurement factor loadings must be significant and above the threshold value of 0.70, (2) construct reliabilities must exceed 0.80, and (3) average variance extracted (AVE) by each construct must exceed the variance due to measurement error for that construct (i.e., AVE should exceed 0.50). As evident from the measurement model in Table 10, the loadings of the measurement items on their respective factors were above the threshold value of 0.70 and all were significant ($p < 0.05$). Composite reliabilities of constructs and Cronbach's alpha ranged between 0.86 and 0.99 and AVE ranged from 0.77 to 0.92. Thus, all three scales met the norms for convergent validity.

In addition, for satisfactory discriminant validity, the square root of AVE from the construct should be greater than the variance shared between the construct and other constructs in the model [32, 81]. As seen from the factor correlation matrix in Table 11, all of the square roots of AVE exceeded interconstruct correlations, providing strong evidence of discriminant validity.

From the beginning of the instrument development process, we envisioned each of the six individual SaaS-Qual factors as reflective latent constructs, while SaaS-Qual itself would be a formative second-order latent construct. Our decision to use the reflective indicator specification for the six SaaS-Qual factors is consistent with several key criteria recommended by Jarvis et al. [41] for choosing that specification over the formative indicator specification: the relative homogeneity and hence interchangeability of scale items within each factor, the high degree of covariation among items within each factor, and the expectation that indicators within each factor (e.g., Features) are likely to be affected by the same antecedents (e.g., Web site design characteristics) and have similar consequences (e.g., increase or decrease in transaction speed or usability).

On the other hand, and in line with previous studies [13, 14], we chose to model the SaaS-Qual construct as a formative (i.e., aggregate) second-order construct consisting of a weighted, linear combination of the six SaaS-Qual factors [41, 67]. In other words, each of the six factors of SaaS-Qual is assumed to be a distinct construct capable of varying independently from others. For example, a SaaS-based application might be highly secure, but not be responsive, thus making the integration of the application into day-to-day work more difficult. This independent nature of the six factors calls for a formative model in which the six dimensions are not treated as reflections of a single underlying overall construct (i.e., SaaS-Qual), but instead the overall construct is seen as "produced/formed by" the combination of the six underlying constructs [4, 13, 67].

Table 10. Assessment of Reflective Measurement Models: Loadings and Reliability

Constructs	Number of indicators	Range of loadings[1]	Cronbach's alpha	Composite reliability (p_c)	AVE
Rapport[2]	9	0.79–0.96	0.93	0.95	0.83
Responsiveness[2]	9	0.86–0.97	0.98	0.99	0.87
Reliability[2]	5	0.81–0.96	0.92	0.93	0.83
Flexibility[2]	6	0.87–0.96	0.96	0.97	0.86
Features[2]	7	0.94–0.97	0.95	0.98	0.92
Security[2]	6	0.70–0.95	0.97	0.98	0.77
SaaS continuance intention	3	0.84–0.87	0.86	0.89	0.73
Satisfaction	4	0.94–0.97	0.97	0.98	0.92
Perceived usefulness	4	0.87–0.92	0.92	0.94	0.81

[1] All factor loadings are significant at least at the $p < 0.05$ level.

[2] All SaaS-Qual constructs were measured based on difference scores between perceived and minimum accepted service levels.

Table 11. Correlation Matrix

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

Notes: Diagonal elements (in boldface) 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 significant.

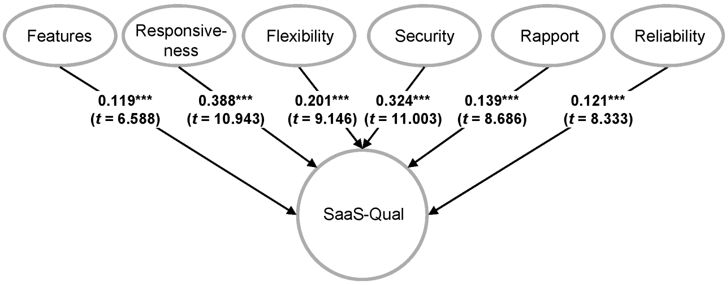

Figure 2. Second-Order Construct Results

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; $n = 172$.

To estimate a formative second-order model of SaaS-Qual (see Figure 2), we modeled the coefficients of each first-order factor to the second-order factor (i.e., Type II second-order model [29]) using a principal components factor analysis, following the procedure in Diamantopoulos and Winklhofer [28]. In this procedure, principal component scores are first computed for the first-order latent variables and then entered as formative indicators for the second-order latent variable. The assessment of SaaS-Qual as second-order factor involved examining the correlations among the first-order factors. Tanaka and Huba [87] argue for the possible validity of a second-order factor, if the first-order factors are highly correlated. Table 11 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. 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 [5], demonstrating that the content captured by the first-order factors are distinct from one another and indicative of discriminant validity. The coefficients (β-values or weights) of the first-order enabling factors to the second-order factors are statistically significant, providing justification for the existence of the hypothesized formative second-order model [17, 30].

Following mediation tests using PLS [38, 40], we also tested whether the second-order construct of overall SaaS-Qual fully mediated the impact of the first-order factors on an important criterion variable (i.e., in our study, 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 variables 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 satisfaction. This supports the conceptualization of overall SaaS-Qual as second-order construct.

Before proceeding with the nomological validity test, we analyzed the six SaaS-Qual factors for possible multicollinearity. Multicollinearity is a concern for formative constructs given that multiple indicators/aggregate constructs are jointly predicting a latent construct in analogous fashion to variables in a multiple regression [67], which can lead to unstable indicator weights [13]. Although PLS is reasonably robust against multicollinearity and skewed responses [12], nonetheless, we performed the relevant assessment. None of the bivariate correlations between the 6 SaaS-Qual dimensions were above 0.90, a potential indication of collinearity [86]. Furthermore, the maximum variance inflation factor of the 6 SaaS-Qual dimensions was 2.54, and so below the suggested tolerance range of 3 to 10 [35].

Nomological/Predictive Validity Test of the SaaS-Qual Scale

Nomological validity is the ability of a new measure to perform as expected in a network of known causal relations and well-established measures [82]. In order to test the nomological validity of SaaS-Qual, we placed it within a nomological network that is rooted in the research area of IS continuance and more specifically in the work of Bhattacherjee [9]. IS continuance, IS continuance behavior, or IS continuous usage describe behavioral patterns reflecting continued use of a particular IS. Continuance refers to a form of post-adoption behavior. Although the term *post-adoption* actually refers to a set of behaviors that follow initial acceptance [76], in the literature it is often used as a synonym for continuance [44]. In this study, we limit ourselves to the terms *IS continuance* or *continued IS usage behavior.*

An influential model on post-acceptance of IS continuance was developed by Bhattacherjee [9], who borrowed heavily from expectation–confirmation theory (ECT), a theory with explicit focus on a user's psychological motivations that emerge after initial adoption, and developed an IS continuance model that includes concepts such as satisfaction, confirmation, and usefulness as the main antecedents of intention to IS continuance behavior. Based on ECT, Bhattacherjee [9] demonstrated that an IS user's continuance decisions are similar to a consumer's repurchase decisions (as they also follow an initial adoption phase), are influenced by the initial use, and can possibly 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 [26, 39, 53] 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. Bhattacherjee's and also follow-up models used only few and rather abstract items to measure how well a user's expectations toward a system's performance and service levels have been met.

However, delivering actionable guidance to practitioners for specific application settings (such as SaaS) about service quality–related causes of customers' discontinuance intentions requires a more fine-grained conceptualization of service quality confirmation. This provided us with a double opportunity. First, Bhattacherjee's IS continuance model can serve as our nomological network, where our newly developed measure of SaaS-Qual can stand in for Bhattacherjee's variable of service quality confirmation. Second, assuming that SaaS-Qual exhibits sufficient nomological validity, it can replace Bhattacherjee's original and rather abstract construct and thus provide a more robust and more highly conceptualized variable within the IS continuance model

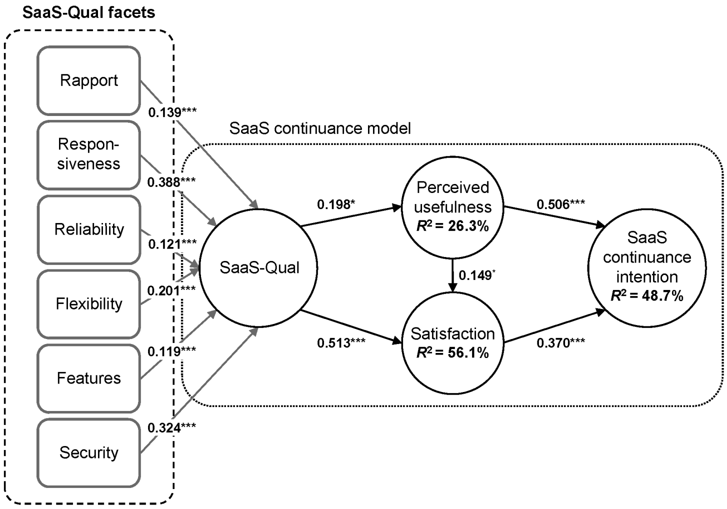

Figure 3. Structural Model Results

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; $n = 172$.

for future researchers in the area. The full model, which serves as our nomological network, is seen in Figure 3.

SmartPLS 2.0 [75] was used to test the relationships between SaaS-Qual and the other constructs (see Table 8) in the nomological network. Overall, the hypothesized research model was supported, indicating strong nomological validity for SaaS-Qual. First, as seen by the structural model results presented in Figure 3, the coefficients are in the appropriate direction and are all statistically significant (a bootstrapping resampling procedure was performed to obtain estimates of standard errors for testing the statistical significance of path coefficients using *t*-tests [16]). Second, the predictive power of the model is good. Overall, 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 ($R^2 = 0.263$). In an alternative structural model, we also tested the direct link between SaaS-Qual and SaaS continuance intention. Consistent with previous studies [9, 52, 53], SaaS-Qual did not have a significant effect on SaaS continuance intention, suggesting that the effect of SaaS-Qual is fully mediated by perceived usefulness and satisfaction. Third, *Responsiveness* ($\beta = 0.388$; $p < 0.001$) and *Security* ($\beta = 0.324$; $p < 0.001$) are the strongest factors to contribute to SaaS-Qual's impact on satisfaction and perceived usefulness.

The predictive relevance of SaaS-Qual was also tested by means of the nonparametric Stone–Geisser test [15] that is based on blindfolding procedures that systematically assumes that a part of the raw data is missing during the parameter estimation. The obtained parameter estimates are then used to reproduce the raw data matrix. The Stone–Geisser test thus shows how well the data collected empirically can be reconstructed with the help of the model and the PLS parameters [31]. The Stone–Geisser

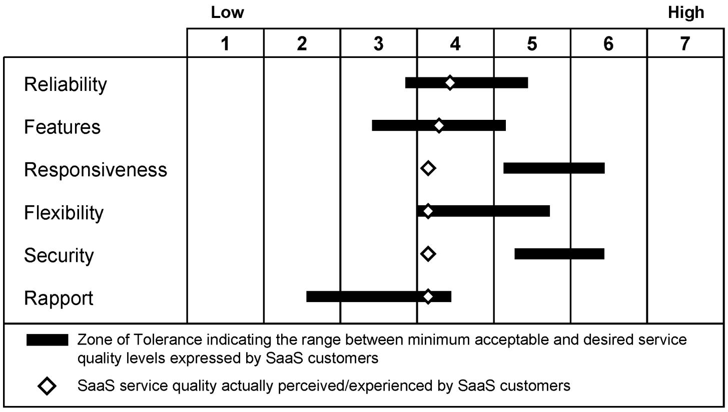

Figure 4. Zones of Tolerance Assessments of SaaS Service Quality (n = 172)

Q^2 value for SaaS-Qual lies well above the threshold level of zero (i.e., $Q^2 = 0.466$), thus validating SaaS-Qual's predictive relevance [36].

Prescriptive Validity of the SaaS-Qual Scale

Beyond verifying the nomological validity of SaaS-Qual, we wanted to use the research model to also test the level at which SaaS customers in our sample felt that their service quality expectations were being met. Beyond the inherent interest in such results, we believe that they can also serve as a "proof of concept" for SaaS-Qual, providing an example of how the instrument can be used empirically by researchers or practitioners alike. Therefore, we analyzed the mean values of each of the six factors regarding their perceived service quality level and how they were positioned relative to their ZOTs (see Figure 4).

All factors met the ZOT except for two: *Responsiveness* ($\mu_{minimum} = 5.22$; $\mu_{desired} = 6.49$; $\mu_{performance} = 4.24$) and *Security* ($\mu_{min} = 5.29$; $\mu_{des} = 6.49$; $\mu_{per} = 4.22$) were far below the minimum acceptable service quality level. Interestingly, these are exactly the two factors that had the most significant influence on customer satisfaction, the highest values for minimum acceptable expectations, and also the smallest ZOTs. While the perceived performance of *Reliability* ($\mu_{min} = 3.73$; $\mu_{des} = 5.53$; $\mu_{per} = 4.43$) and *Flexibility* ($\mu_{min} = 4.03$; $\mu_{des} = 5.77$; $\mu_{per} = 4.26$) was at the lower end of their ZOTs, *Rapport* ($\mu_{min} = 2.69$; $\mu_{des} = 4.56$; $\mu_{per} = 4.15$) was at the upper end. Last, *Features* ($\mu_{min} = 3.43$; $\mu_{des} = 5.18$; $\mu_{per} = 4.37$) hit the ZOT right at its center.

Discussion

Theoretical and Practical Contributions

In a recent article, Bardhan et al. [6] laid out a framework for new research in service science, emphasizing that the increasing importance of IT services has opened a vast arena of research opportunities in this area for IS researchers. In this paper, we reported on the results of a study that addresses one of the research opportunities Bardhan et al. [6] propose: experimental and behavioral research concentrating on IT service users. More specifically, we developed a service quality measure for SaaS clients, which we call "SaaS-Qual," by adapting prior service quality measures related to SERVQUAL and its variations. We conducted thorough validity tests of SaaS-Qual, including a test of its criterion-related validity through a nomological network based on IS continuance and expectation confirmation theories. Our study provides several important theoretical as well as practical implications.

The primary research contribution lies in the development of a context-specific service quality measure. There is a recent tradition in the IS literature of the development of such measures, beginning with early appropriations of SERVQUAL for measuring service quality of IS in general [46, 90] to later attempts in creating context-specific service quality measures, primarily for general online services [33, 85, 96]. Most related to our work was the development of ASP-Qual, which measured service quality of ASP-hosted online Web stores [79]. However, as we discussed extensively in this paper, SaaS presents unique service quality challenges for SaaS vendors and therefore unique service quality expectations for their clients. We believe that SaaS-Qual provides a theoretically sound measurement instrument that addresses those distinctive characteristics of SaaS. It should thus be able to support a wide range of IS studies as researchers attempt to understand a variety of important issues related to the SaaS space. Although individual researchers may focus on somewhat different issues arising in the SaaS or cloud context, and may need to develop or refine measures for their own specialized scales, the existence of a validated baseline measure should be a major advantage. It provides a good starting point that will speed up and enhance the quality of measurement for many studies.

Our study also provides theoretical contributions to the area of IS continuance, since we used an IS continuance model as the nomological network for testing SaaS-Qual's predictive validity. An important aspect of our study is the replacement of the original confirmation measure (taken from expectation confirmation theory) in the IS continuance model with SaaS-Qual. Our measure provides a more detailed, more granular approach to understanding whether expectations are confirmed or disconfirmed, than the original three-item scale used. Overall, with this more fine-grained conceptualization of service quality confirmation, the variance explained of customer satisfaction is increased to 0.56, almost double of the variance explained in Bhattacherjee's [9] initial model (0.33). Further, the variance explained of the other two dependent variables is also increased considerably. By developing SaaS-Qual into a formative second-order construct with six reflective first-order constructs, we have provided a multidimensional measure, where each dimension is independent of the others. This approach could be very useful in future studies that use expectation confirmation theory, especially where confirmation of service quality is in question.

The results from our test of the nomological network provide interesting insights on the role of SaaS service quality within the IS continuance model. While in the original model Bhattacherjee [9] found that expectation confirmation had more or less equal impact on perceived usefulness and satisfaction of online banking by its users, we found that confirmation of SaaS service quality had a much larger impact (more than threefold) on satisfaction than it did on perceived usefulness. In other words, meeting SaaS service quality expectations leads to a strong overall feeling of satisfaction with the system, but is not seen as increasing the productivity and efficiency of the clients' operations to an equal degree. One possible explanation is that given the relatively young age of SaaS systems, clients have not had enough time to fully evaluate their usefulness. Follow-up longitudinal studies could provide interesting insights on this subject.

Our results also shed some light on the specific nature of service quality expectations for SaaS clients. By using the ZOT approach, we were able to identify the specific areas in which SaaS clients in our sample felt that their expectations had not been met. First, as key SaaS-Qual factors driving the influence on customer satisfaction and perceived usefulness, we found that *Responsiveness* and *Security* have the strongest impact. Second, we found that all factors were meeting their ZOTs except for those that had the strongest influence on customer satisfaction and perceived usefulness, the highest values for minimum acceptable expectations, and also the narrowest ZOTs. *Responsiveness* and *Security* 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. In general, we believe that using the SaaS-Qual instrument with the ZOT approach can provide both researchers and practitioners with a clear picture of where corrective action is necessary to improve service quality for SaaS users.

IT managers responsible for selecting or renewing SaaS-based solutions can learn from the results of our 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. According to our sample, IT managers should particularly focus on a SaaS vendor's operations management capabilities in the areas of *Responsiveness* and *Security.* More specifically, they could negotiate contractual uptime guarantees or 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 service-level agreements (SLAs) on security, 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., key performance indicator metrics), and enforcement plans (e.g., penalty/reward definitions).

Limitations and Future Research

Our research study is not without limitations. First, we collected data from a single respondent within the organization. Given the nature of the survey items, the majority of respondents are IT 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. Nonetheless, further confirmatory research needs to be done with broad samples of IT executives/managers, professionals, and users to increase the applicability of the SaaS-Qual instrument in organizations [42]. Furthermore, given that SaaS-Qual has completed initial development and refinement, it is now appropriate to investigate some other important areas, such as business-to-consumer (B2C) or business-to-government (B2G) SaaS solutions.

Second, 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 SaaS-Qual. While 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 [22]), other scholars (e.g., [47]) have emphasized the diagnostic and thus practical value of the difference score-based (IS) service quality measurements. Although our study demonstrated satisfying construct validity and predictive values of SaaS-Qual, future studies that may want to use a differently conceptualized measurement instrument could replicate this study with alternative measurement models (e.g., with perceptions-only items) and compare their results to SaaS-Qual. Also, other analytical techniques could be used (e.g., polynomial regression and response surface modeling) to investigate the impact of interactions between expected service levels and perceived service performance (i.e., pre-exposure expectation and post-exposure experience as two separate component scores) on outcome variables such as satisfaction or perceived usefulness [48, 88].

Conclusion

Given a growing service orientation in the IS industry and with SaaS-based software delivery quickly gaining importance, it has become critical for companies to regularly assess the service quality factors of SaaS services and their importance for continued IS usage. At the same time, it has presented both challenges and opportunities for IS researchers. With this study, we provided both a service quality measurement instrument with practical significance and applications, as well as some important theoretical contributions in IS continuance research and service science in general. We hope that it will serve as a springboard for future research studies and also aid SaaS vendors in better managing their clients' expectations. To the extent that researchers converge on standard measures, SaaS-Qual may also serve as a validated baseline measure that makes it much easier to compare and consolidate findings across studies.

Notes

1. The hit ratio is a measure of how well the items tap into the constructs and is calculated as the ratio of "correct" item placements to total placements across all dimensions [59]. There are

no strict guidelines for the hit ratio, but a minimum of 75 percent is generally acceptable [37]. Kappa assesses agreement between judges and the commonly accepted threshold for kappa is 0.70 [49].

2. A covariance-based structural equation modeling application (such as LISREL or Amos) could not be used to conduct a confirmatory factor analysis, because our sample size (n = 172) was below the recommended threshold of 5 to 10 times as many observations as there are items in the research model [35].

References

1. Anderson, J.C., and Gerbing, D.W. Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103,* 3 (1988), 411–423.

2. Armbrust, M.; Fox, A.; Griffith, R.; Joseph, A.D.; Katz, R.; Konwinski, A.; Lee, G.; Patterson, D.; Rabkin, A.; Stoica, I.; and Zaharia, M. A view of cloud computing. *Communications of the ACM, 53,* 4 (2010), 50–58.

3. Bagozzi, R.P. An examination of the validity of two models of attitude. *Multivariate Behavioral Research, 16,* 3 (1981), 323–359.

4. Bagozzi, R.P., and Edwards, J.R. A general approach for representing constructs in organizational research. *Organizational Research Methods, 1,* 1 (1998), 45–87.

5. Bagozzi, R.P.; Yi, Y.; and Phillips, L.W. Assessing construct validity in organizational research. *Administrative Science Quarterly, 36,* 3 (1991), 421–458.

6. Bardhan, I.R.; Demirkan, H.; Kannan, P.K.; Kauffman, R.J.; and Sougstad, R. An interdisciplinary perspective on IT services management and service science. *Journal of Management Information Systems, 26,* 4 (Spring 2010), 13–64.

7. Barnes, S.J., and Vidgen, R. An evaluation of cyber-bookshops: The WebQual method. *International Journal of Electronic Commerce, 6,* 1 (2001) 11–30.

8. Benlian, A.; Hess, T.; and Buxmann, P. Drivers of SaaS-adoption: An empirical study of different application types. *Business & Information Systems Engineering, 1,* 5 (2009), 357–369.

9. Bhattacherjee, A. Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly, 25,* 3 (2001), 351–370.

10. Bollen, K.A., and Jackman, R.W. Regression diagnostics: An expository treatment of outliers and influential cases. In J. Fox and J.S. Long (eds.), *Modern Methods of Data Analysis.* Newbury Park, CA: Sage, 1990, pp. 257–291.

11. Byrne, B.M. *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications, and Programming.* Mahwah, NJ: Psychology Press, 1998.

12. Cassel, C.; Hackl, P.; and Westlund, A.H. Robustness of partial least-squares method for estimating latent variable quality structures. *Journal of Applied Statistics, 26,* 4 (1999), 435–446.

13. Cenfetelli, R.T., and Bassellier, G. Interpretation of formative measurement in information systems research. *MIS Quarterly, 33,* 4 (2009), 689–707.

14. Cenfetelli, R.T.; Benbasat, I.; and Al-Natour, S. 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 (2008), 161–181.

15. Chin, W.W. Issues and opinion on structural equation modelling. *MIS Quarterly, 22,* 1 (1998), vii–xvi.

16. Chin, W.W. The partial least squares approach for structural equation modelling. In G.A. Marcoulides (ed.), *Modern Methods for Business Research.* Hillsdale, NJ: Lawrence Erlbaum, 1998, pp. 295–336.

17. Chin, W.W.; Marcolin, B.L.; and Newsted, P.R. 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 (2003), 189–217.

18. Chou, T. *Seven Business Models for Software.* San Mateo, CA: Active Book Press, 2008.

19. Choudhary, V. Comparison of software quality under perpetual licensing and software as a service. *Journal of Management Information Systems, 24,* 2 (Fall 2007), 141–165.

20. Churchill, G.A. A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16,* 1 (1979), 64–73.

21. Cohen, J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 1 (1960), 37–46.

22. Cronin, J.J., Jr., and Taylor, S.A. SERVPERF versus SERVQUAL: Reconciling performance-based and perceptions-minus-expectations measurement of service quality. *Journal of Marketing, 58,* 1 (1994), 125–131.

23. Cusumano, M.A. Cloud computing and SaaS as new computing platforms. *Communications of the ACM, 53,* 4 (2010), 27–29.

24. Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13,* 3 (1989), 319–339.

25. Davis, F.D.; Bagozzi, R.P.; and Warshaw, P.R. User acceptance of computer technology: A comparison of two theoretical models. *Management Science, 35,* 8 (1989), 982–1003.

26. Deng, L.; Turner, D.E.; Gehling, R.; and Prince, B. User experience, satisfaction, and continual usage intention of IT. *European Journal of Information Systems, 19,* 1 (2010), 60–75.

27. DeVellis, R.F. *Scale Development—Theory and Applications.* Thousand Oaks, CA: Sage, 2003.

28. Diamantopoulos, A., and Winklhofer, H.M. Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 38,* 2 (2001), 269–277.

29. Diamantopoulos, A.; Riefler, P.; and Roth, K.P. Advancing formative measurement models. *Journal of Business Research, 61,* 12 (2008), 1203–1218.

30. Edwards, J.R. Multidimensional constructs in organizational behavior research: An integrative analytical framework. *Organizational Research Methods, 4,* 2 (2001), 144–192.

31. Fornell, C., and Cha, J. Partial least squares. In R.P. Bagozzi (ed.), *Advanced Methods of Marketing Research.* Cambridge, MA: Blackwell, 1994, pp. 52–78.

32. Fornell, C., and Larcker, D.F. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18,* 1 (1981), 39–50.

33. Gefen, D. Customer loyalty in e-commerce. *Journal of the Association for Information Systems, 3,* 1 (2002), 27–51.

34. Grover, V.; Myun Joong, C.; and Teng, J.T.C. The effect of service quality and partnership on the outsourcing of information systems functions. *Journal of Management Information Systems, 12,* 4 (Spring 1996), 89–116.

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

36. Henseler, J.; Ringle, C.M.; and Sinkovics, R.R. The use of partial least squares path modeling in international marketing. *Advances in International Marketing, 20* (2009), 277–319.

37. Hinkin, T.R. A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1,* 1 (1998), 104–121.

38. Hoyle, R.H., and Kenny, D.A. Sample size, reliability, and tests of statistical mediation. In R.H. Hoyle (ed.), *Statistical Strategies for Small Sample Research.* Thousand Oaks, CA: Sage, 1999, pp. 197–223.

39. Hsu, M.; Chiu, C.; and Ju, T. Determinants of continued use of the WWW: An integration of two theoretical models. *Industrial Management & Data Systems, 104,* 9 (2004), 766–775.

40. Iacobucci, D.; Saldanha, N.; and Deng, X. A meditation on mediation: Evidence that structural equations models perform better than regressions. *Journal of Consumer Psychology, 17,* 2 (2007), 139–153.

41. Jarvis, C.B.; MacKenzie, S.B.; Podsakoff, P.M.; Mick, D.G.; and Bearden, W.O. A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research, 30,* 2 (2003), 199–218.

42. Jiang, J.J.; Klein, G.; and Carr, C.L. Measuring information system service quality: SERVQUAL from the other side. *MIS Quarterly, 26,* 2 (2002), 145–166.

43. Jöreskog, K.G., and Sörbom, A.S. *LISREL 8: User's Reference Guide.* Chicago: Scientific Software International, 1996.

44. Karahanna, E.; Straub, D.W.; and Chervany, N.L. Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs. *MIS Quarterly, 23,* 2 (1999), 183–213.

45. Kettinger, W.J., and Lee, C.C. Perceived service quality and user satisfaction with the information services function. *Decision Sciences, 25,* 5–6 (1994), 737–766.

46. Kettinger, W.J., and Lee, C.C. Pragmatic perspectives on the measurement of information systems service quality. *MIS Quarterly, 21,* 2 (1997), 223–240.

47. Kettinger, W.J., and Lee, C.C. Zones of tolerance: Alternative scales for measuring information systems service quality. *MIS Quarterly, 29,* 4 (2005), 607–623.

48. Klein, G.; Jiang, J.J.; and Cheney, P.H. Resolving difference score issues in information systems research. *MIS Quarterly, 33,* 4 (2009), 811–826.

49. Krippendorff, K. *Content Analysis: An Introduction to Its Methodology.* Beverly Hills, CA: Sage, 2004.

50. Kumar, N.; Stern, L.W.; and Anderson, J.C. Conducting interorganizational research using key informants. *Academy of Management Journal, 36,* 6 (1993), 1633–1651.

51. Lewis, R.C., and Booms, B.H. The marketing aspects of service quality. In L.L. Berry, G.L. Shostack, and G. Upah (eds.), *Emerging Perspectives on Services Marketing.* Chicago: American Marketing Association, 1983, pp. 99–107.

52. Limayem, M., and Cheung, C.M.K. Understanding information systems continuance: The case of Internet-based learning technologies. *Information & Management, 45,* 4 (2008), 227–232.

53. Limayem, M.; Hirt, S.G.; and Cheung, C.M.K. How habit limits the predictive power of intention: The case of information systems continuance. *MIS Quarterly, 31,* 4 (2007), 705–737.

54. Loiacono, E.T.; Watson, R.T.; and Goodhue, D.L. WebQual: An instrument for consumer evaluation of Web sites. *International Journal of Electronic Commerce, 11,* 3 (2007), 51–87.

55. Ma, Q.; Pearson, J.M.; and Tadisina, S. An exploratory study into factors of service quality for application service providers. *Information & Management, 42,* 8 (2005), 1067–1080.

56. Mather, T.; Kumaraswamy, S.; and Latif, S. *Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance.* Sebastopol, CA: O'Reilly Media, 2009.

57. Menken, I. *SaaS—The Complete Cornerstone Guide to Software as a Service Best Practices: Concepts, Terms, and Techniques for Successfully Planning, Implementing and Managing SaaS Solutions.* London: Emereo, 2008.

58. Mertz, S.A.; Eschinger, C.; Eid, T.; Huang, H.H.; Pang, C.; and Pring, B. Market trends: Software as a service, worldwide, 2008–2013. Gartner, Stamford, CT, 2009.

59. Moore, G.C., and Benbasat, I. Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research, 2,* 3 (1991), 192–222.

60. Mulholland, J.A.; Pyke, J.; and Fingar, P. *Enterprise Cloud Computing: A Strategy Guide for Business and Technology Leaders.* Tampa, FL: Meghan-Kiffer Press, 2010.

61. Parasuraman, A.; Berry, L.L.; and Zeithaml, V.A. Refinement and reassessment of the SERVQUAL scale. *Journal of Retailing, 67,* 4 (1991), 420–450.

62. Parasuraman, A.; Zeithaml, V.A.; and Berry, L.L. A conceptual model of service quality and its implications for future research. *Journal of Marketing, 49,* 4 (1985), 41–50.

63. Parasuraman, A.; Zeithaml, V.A.; and Berry, L.L. SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing, 64,* 1 (1988), 12–40.

64. Parasuraman, A.; Zeithaml, V.A.; and Berry, L.L. Reassessment of expectations as a comparison standard in measuring service quality: Implications for further research. *Journal of Marketing, 58,* 1 (1994), 111–124.

65. Parasuraman, A.; Zeithaml, V.A.; and Malhotra, A. E-S-QUAL: A multiple-item scale for assessing electronic service quality. *Journal of Service Research, 7,* 3 (2005), 213–233.

66. Pavlou, P.A., and El Sawy, O.A. From IT leveraging competence to competitive advantage in turbulent environments: The case of new product development. *Information Systems Research, 17,* 3 (2006), 198–227.

67. Petter, S.; Straub, D.; and Rai, A. Specifying formative constructs in information systems research. *MIS Quarterly, 31,* 4 (2007), 623–656.

68. Pettey, C. Gartner says 25 percent of new business software will be delivered as software as a service by 2011. Gartner, Stamford, CT, 2006.

69. Pitt, L.F.; Watson, R.T.; and Kavan, C.B. Service quality: A measure of information systems effectiveness. *MIS Quarterly, 19,* 2 (1995), 173–187.

70. Pitt, L.F.; Watson, R.T.; and Kavan, C.B. Measuring information systems service quality: Concerns for a complete canvas. *MIS Quarterly, 21,* 2 (1997), 209–221.

71. Podsakoff, P.M., and Organ, D.W. Self-reports in organizational research: Problems and prospects. *Journal of Management, 12,* 4 (1986), 531–544.

72. Pring, B., and Lo, T. Dataquest insight: SaaS adoption trends in the U.S. and U.K. Gartner, Stamford, CT, 2009.

73. Reese, G. *Cloud Application Architectures: Building Applications and Infrastructure in the Cloud.* Sebastopol, CA: O'Reilly Media, 2009.

74. Richardson, H.A.; Simmering, M.J.; and Sturman, M.C. A tale of three perspectives: Examining post hoc statistical techniques for detection and correction of common method variance. *Organizational Research Methods, 12,* 4 (2009), 762–800.

75. Ringle, C.M.; Wende, S.; and Will, A. SmartPLS 2.0 (M3) beta. University of Hamburg, 2005.

76. Rogers, E.M. *Diffusion of Innovations.* New York: Free Press, 1962.

77. Rossiter, J.R. The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing, 19,* 4 (2002), 305–335.

78. Segars, A., and Grover, V. Re-examining perceived ease of use and usefulness: A confirmatory factor analysis. *MIS Quarterly, 17,* 4 (1993), 517–525.

79. Sigala, M. The ASP-Qual model: Measuring ASP service quality in Greece. *Managing Service Quality, 14,* 1 (2004), 103–114.

80. Stewart, D.W., and Shamdasani, P.N. *Focus Groups: Theory and Practice.* Newbury Park, CA: Sage, 2004.

81. Straub, D. Validating instruments in MIS research. *MIS Quarterly, 13,* 2 (1989), 147–169.

82. Straub, D.; Boudreau, M.-C.; and Gefen, D. Validation guidelines for IS positivist research. *Communications of the Association for Information Systems, 13,* 24 (2004), 380–427.

83. Susarla, A.; Barua, A.; and Whinston, A.B. A transaction cost perspective of the "software as a service" business model. *Journal of Management Information Systems, 26,* 2 (Fall 2009), 205–240.

84. Susarla, A.; Barua, A.; and Whinston, A.B. Multitask agency, modular architecture, and task disaggregation in SaaS. *Journal of Management Information Systems, 26,* 4 (Spring 2010), 87–117.

85. Swaid, S.I., and Wigand, R.T. Measuring the quality of e-service: Scale development and initial validation. *Journal of Electronic Commerce Research, 10,* 1 (2009), 13–28.

86. Tabachnick, B.G., and Fidell, L.S. *Using Multivariate Statistics,* 4th ed. New York: HarperCollins, 2000.

87. Tanaka, J.S., and Huba, G.J. Confirmatory hierarchical factor analysis of psychological distress measures. *Journal of Personality and Social Psychology, 46,* 3 (1984), 621–635.

88. Venkatesh, V., and Goyal, S. Expectation disconfirmation and technology adoption: Polynomial modeling and response surface analysis. *MIS Quarterly, 34,* 2 (2010), 281–303.

89. Waters, B. Software as a service: A look at the customer benefits. *Journal of Digital Asset Management, 1,* 1 (2005), 32–39.

90. Watson, R.T.; Pitt, L.F.; and Kavan, C.B. Measuring information systems service quality: Lessons from two longitudinal case studies. *MIS Quarterly, 22,* 1 (1998) 61–79.

91. Weier, M.H. Google, workday outages show SaaS isn't perfect. *InformationWeek* (October 21, 2009) (available at www.informationweek.com/news/services/saas/220800123/).

92. Werts, C.E.; Linn, R.L.; and Joreskog, K.G. Intraclass reliability estimates: Testing structural assumptions. *Educational and Psychological Measurement, 34,* 1 (1974), 25–33.

93. Wolfinbarger, M., and Gilly, M.C. eTailQ: Dimensionalizing, measuring, and predicting etail quality. *Journal of Retailing, 79,* 3 (2003), 183–198.

94. Xin, M., and Levina, N. Software-as-a-service model: Elaborating client-side adoption factors. Paper presented at the Twenty-Ninth International Conference on Information Systems, Paris, December 15–17, 2008.

95. Zeithaml, V.A.; Berry, L.L.; and Parasuraman, A. The nature and determinants of customer expectations of service. *Journal of the Academy of Marketing Science, 21,* 1 (1993), 1–12.
96. Zeithaml, V.A.; Parasuraman, A.; and Malhotra, A. Service quality delivery through Web sites: A critical review of extant knowledge. *Journal of the Academy of Marketing Science, 30,* 4 (2002), 362–375.

Copyright of Journal of Management Information Systems is the property of M.E. Sharpe Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.