Towards a Framework for the Evaluation Design of Enterprise Social Software

Completed Research Paper

Christian Herzog
University of Osnabrück
Osnabrück, Germany
christian.herzog@uni-osnabrueck.de

Alexander Richter
University of Zurich
Zürich, Switzerland
arichter@ifi.uzh.ch

Melanie Steinhüser
University of Osnabrück
Osnabrück, Germany
melanie.steinhüser@uni-osnabrueck.de

Abstract

While the use of Enterprise Social Software (ESS) increases, reports from science and practice show that evaluating its impact remains a major challenge. Various interests and points of view make each ESS evaluation an individual matter and lead to diverse requirements. In this paper, we propose a design theory that highlights the various design options and ensures completeness and consistency. Based on a comprehensive literature analysis, as well as an interview study with 31 ESS experts from 29 companies, we suggest a conceptual framework intended as decision support for the ESS evaluation design for different stakeholders. Beyond providing an orientation the framework also reveals six evaluation classes that represent typical application instantiations and can be understood as principles of implementation. A first validation in five organizations confirms that the framework can lead to a more efficient evaluation design and to increased effectiveness during the evaluation phase.

Keywords: Enterprise Social Software, Evaluation, Design Theory, Framework

Motivation

During the past years many organizations have adopted Enterprise Social Software (ESS) to support organizational communication, knowledge and innovation management (Aral et al. 2013; Faraj et al. 2011; Kane et al. 2009). While the capacity to use these tools is slowly maturing (Kiron et al. 2013), companies are increasingly realizing benefits and competitive advantages from using ESS, such as improved productivity, better knowledge sharing, and enhanced employee innovativeness (Chui et al. 2012; Gray et al. 2011). Nevertheless, organizations still face the question how to comprehensively evaluate these impacts in their particular cases. As a result, researchers and practitioners report a lack of knowledge about ESS success measurement (Kraut and Resnick 2012; Kügler et al. 2013; McAfee 2009). With this paper we aim at contributing to the understanding of ESS evaluation by developing a framework that supports the ESS evaluation design for different stakeholders.

The why and how of evaluating the impact not only of ESS but of any type of Information System (IS) is a major challenge. A multitude of procedures and models have been developed (DeLone and McLean 1992; Gable et al. 2003), and each model has its own approach to different systems, stakeholder perspectives, or theories (Kronbichler et al. 2010). However, while useful for specific purposes, these models' contextual references do not help any company design its individual evaluation approach. As a prominent example, DeLone and McLean's IS Success Model does not recognize that different stakeholders in an organization may come to different conclusions about the success of the same IS. At the same time, there is a lack of studies that provide an overview of factors to be considered and thus provide guidelines on how IS can be evaluated. In one of these rare studies, Grover et al. (1996) use the dimensions of evaluative referent, unit of analysis, and evaluation objective to both unify and distinguish criteria associated with IS effectiveness. Similarly, Seddon et al. (1999) develop measurement guidelines based on questions gleaned from organizational psychology (Cameron and Whetten 1983) that are relevant to IT professionals seeking to measure IS effectiveness. These questions concern perspective, domain, and level of analysis, as well as purpose, time, data types, and referents. The above mentioned studies provide valuable findings and therefore, constitute a useful starting point for our aim to develop a framework for the design of an ESS evaluation. Nevertheless, a number of studies have shown that ESS has various characteristics which have to be considered when designing an evaluation approach. E.g. ESS helps users to easily create own content (Du and Wagner, 2006) and leads to a higher degree of networking of the employees (DiMicco et al. 2008; Richter and Riemer 2009). Moreover, ESS is mostly used to support unstructured tasks and can be characterized as malleable end-user software with no clearly a priori defined usage scenarios. As a consequence, the evaluation of the usage and the resulting benefits is difficult due to the lack of correspondence with one particular business process (Richter and Riemer 2013). These characteristics considerably distinguish ESS from other types of IS which requires adapting established theories, or possibly developing new ones (Majchrzak 2009).

Our study sheds light on the design options, eventually leading to a clear set of guidelines for ESS evaluation. The application of these guidelines will lead to more efficient evaluation design and to increased effectiveness during the evaluation phase. Our research design is informed by design science which recently has received increasing attention and recognition in the IS community (Baskerville and Pries-Heje 2010; Peffers et al. 2012). Whereas its theoretical basis is still subject of intense discussions (Hooker 2004), the need for and importance of design science for the IS community were recently indicated in several calls for action of prominent journals (e.g. Goes 2014). Our paper, which seeks to answer these calls, posits a design theory based on a conceptual framework and provides guidance for designing ESS evaluations. While debates about terminology are ongoing, a shared understanding as well as agreement about the creation of a design theory have been achieved to some extent (Iivari 1983; Markus et al. 2002; Walls et al. 1992; Gregor and Jones 2007; Gregor et al. 2013; Gregor and Hevner 2013). Being aware of these terminology ambiguities, we structure our paper according to Gregor and Jones' (2007) "anatomy of a design theory," comprising eight core components of an information systems design theory.

Accordingly, in the following section, we describe the (1) design theory's *purpose and scope*. Subsequently, based on a literature analysis, we examine the (2) *justificatory knowledge* that directly flows into the framework's creation. Thereafter, we describe the (3) conceptual framework's *constructs and* (4) *principles of form and function*, as well as its dimensions and their characteristics. In addition to the theoretical basis and in order to refine the results, we conducted 31 interviews with ESS experts from

29 companies. These insights from practice led us to empirically validate the framework and to define (5) *testable propositions*. However, recommendations are required for the conversion to an evaluation procedure. This gap is closed by a section on the (6) *principles of implementation and* (7) *expository instantiation*. Beyond this, the interview study allowed us to identify and analyze specific evaluation scenarios in various business situations where typical combinations of different dimensions' characteristics are applied. These combinations led us to build six classes of ESS evaluation that represent main instantiations, and the subsequent presentation of a design theory to three company representatives as well as two research groups allowed for additional verification. We then discuss the (8) *artifact's mutability*. A summary and an outlook on further research conclude the contribution.

Purpose and Scope

Following Gregor and Jones (2007), a design theory's purpose and scope not only form the goal of and the need for a theoretical foundation of ESS evaluations, but also the (meta-)requirements (Walls et al. 1992) and the differentiation of application areas (Dubin 1978).

In our context, the abstract artifact is an ESS evaluation framework that is designed to support individual ESS evaluations in terms of designing and implementing the evaluation approach. This individuality derives from including various views and influence factors for different situations and interests. In each ESS implementation case, the interests of different internal stakeholders, the use cases or objectives vary, which directly affects how it is evaluated. Additionally, there are further challenges arising from the ESS' peculiarities. As meta requirement (Walls et al. 1992), our design theory claims to work for any kind of ESS evaluation. This means that the technical platform or tool, as well as the influence for instance on the user, group, or organization fall within the focus of the evaluation design and not the general business performance independent from a technical support. Additionally, the provided framework is supposed to optimize the ESS evaluation by the conscious consideration of the different dimensions and characteristics of influence. Our subsequent distinction of evaluation instantiations into six main classes seeks to support practitioners and researchers in designing an evaluation for specific situations.

Evaluation generally involves assessment in every sense, and thus distinguishes itself from similar terms such as *success measurement*, which focuses on measuring positive effects instead of the general impact. As a result, we see success measurement as a kind of evaluation. Furthermore, *measurement* implies quantification. In contrast, our understanding of *evaluation* includes quantitative as well as qualitative assessments. Generally, there are different terms for the effects found in the evaluation literature (Schubert and Williams 2013a). This includes *performance* (Melville et al. 2004), *productivity* (Brynjolfsson and Hitt 1993), *(business) value* (Schryen 2010), *benefit* (Schubert et al. 2011), *acceptance* (Venkatesh et al. 2003), *impact* (Gable et al. 2008) or the afore-mentioned *success* (DeLone and McLean 2003). Our study considers various perspectives of an evaluation, which can be found in existing literature, combines exactly these different perspectives as design options in one framework and provides a novel approach in form of a design theory. Thus, for an analysis of the different dimensions in different perspectives, comprehensive terminology is necessary.

Although there are many parallels to the evaluation of other types of IS, this paper focuses on ESS (as an IS type), which includes all social software tools and platforms used for the company internal corporate communication and collaboration and therefore requires skills in the area of computer-supported cooperative work (Stocker et al. 2012). Examples of ESS include internal blogs, microblogs, wikis, social networks, instant messaging, and especially platforms that include several social software functionalities, such as Atlassian Confluence, Jive, Microsoft SharePoint, IBM Connections, etc. Various terms that are used interchangeably, such as social network technologies (Kettles and David 2008), corporate social software (Steinhüser et al. 2011), enterprise collaboration systems (Schubert and Williams 2013b), enterprise social networks (Richter et al. 2013), or enterprise 2.0 (McAfee 2009), to name a few. Compendiously ESS is a web-based technology that support users' contributions of persistent objects to a shared pool and that enable company-wide responses to these objects. ESS comprises functionalities that visualize profile information and link users with one another (Steinhüser et al. 2015). However, the mentioned characteristics provide an abstract interpretation; accordingly, we don't limit our understanding to particular ESS applications, but adopt a broad view that includes all the types covered by these traits. In summary, the research field is limited to *ESS* and *evaluation* (Dubin 1978), even if we assume that cognitions can be transferred to other types of IS.

Justificatory Knowledge

A design theory is based on several micro-theories (Simon 1996) or kernel theories (Walls et al. 1992). These should explain why a design theory works and should thus justify it (Gregor and Jones 2007). Furthermore, this knowledge base is used to construct design science artifacts (Hevner et al. 2004).

Literature Analysis Approach

The theoretical foundation for the structure and the composition of our evaluation framework is based on prior research on IS and ESS evaluation. As ESS being one type of IS, not only studies on ESS but also on IS in general might provide us with valuable knowledge. An extensive literature review, resulting in 51 scientific publications, helped us to gain an understanding of the different views and influences on evaluation. At the same time, we were able to identify and compile the existing dimensions and characteristics of our conceptual evaluation framework (Webster and Watson 2002). Our search included the online databases EBSCO, IEEE Xplore, ACM Digital Library, AIS Electronic Library, and Google Scholar. Owing to the different terminology for an evaluation, we used the following search terms: "evaluation," "success," "performance," "productivity", "business value," "benefits," "impact" and "acceptance." Our final literature selection includes 31 journal articles, 19 conference proceedings, and one book contribution on the evaluation of IS and ESS. The publication dates are between 1989 and 2013 (10 between 1989 and 1999, 26 between 2000 and 2009, and 15 between 2010 and 2013). The consideration of both ESS and other IS publications helped us to obtain a comprehensive overview of various evaluation approaches and hence to adopt insights from other IS to ESS. The variety of different publications with different approaches highlights the evaluation diversity. On the one hand, this confirms the need for a comprehensive procedure; on the other hand, this raises the challenge to unify the theoretical approaches. To this end, we consulted Mayring's (2000) deductive category application, working with predefined and theory-based aspects of analysis, and formed dimensions of ESS evaluation. For this purpose, each of the authors developed an understanding about dimensions and their particular characteristics, as well as their relevance and meaning for ESS evaluations. Finally, after presenting these individual results to the research group, problematic issues were resolved through discussion (Walsham 2006). Concerning the subsequent interview study, the justificatory knowledge plays a bridging role between the dimensions framework and practical experience (Spagnoletti and Tarantino 2013). Furthermore, the analysis of the interviews helps us to validate the results from the literature and to examine the ESS application in practice.

Results

Analyzing the literature and the deductive category application has led us to distinguish eight dimensions that influence ESS evaluation (time, purpose, level of analysis, perspective, evaluation objective, data, data gathering, and context).

The fundamental dimension to be found in the literature refers to the **time** of evaluation. Stockman (2011) identifies three characteristics: *ex ante*, *ongoing*, and *ex post*. Ex ante, potential costs and benefits of an investment can be calculated, or the system requirements can be assessed, in advance (e.g. Desmarais et al. 1997). Because little data exist at that time, ex ante evaluation is a challenging activity that is usually based on qualitative indices or estimates as well as financial cost calculations (e.g. Stefanou 2001). Additionally studies show that the diffusion of ESS in the company is mainly bottom up whereby a comprehensive ex ante evaluation might be skipped in some cases (Richter and Riemer 2013a). In contrast, several data such as usage statistics, costs, etc. have already incurred in the ex post and ongoing evaluation. Ongoing evaluation is used for interim controls (e.g. Muller et al. 2009) or further development of the system (e.g. Davenport et al. 2004). Finally, ex post means at the end of a project, system lifetime, or given timeframe. System design aspects do not play a role at this time. It is therefore always summative (e.g. Sabherwal et al. 2006). Very few evaluation models consider all phases; Markus and Tanis's (2000) enterprise system experience cycle is an exception.

	Ex ante	Ongoing	Ex post
Time	Desmarais et al. 1997; Irani and Love 2001; Markus and Tanis 2000; Schubert and Williams 2011; Stefanou 2001; Tallon et al. 2000	Davenport et al. 2004; Lin and Lee 2006; Majchrzak et al. 2006; Markus and Tanis 2000; Muller et al. 2009	Markus and Tanis 2000; Sabherwal et al.2006; Schubert and Williams 2011; Tallon et al. 2000

Closely related to the time dimension is the **purpose** served by an evaluation. Basically, an evaluation can seek to *understand* processes, functions (e.g. Lin and Lee 2006; Schubert and Williams 2011) and use of a system (e.g. Li and Gupta 2009). Understanding a system's use is particularly relevant for malleable end-user software such as ESS. It is mostly used to support unstructured tasks, and there are no clearly a priori defined usage scenarios, so the tools can be used for a variety of practices in everyday work (Richter and Riemer 2013a). The purpose *control* is served when comparing the current state with predefined targets (e.g. Rosemann and Wiese 1999). An evaluation is further used to identify requirements and improvement potentials for the further *development* of a system (e.g. Davenport et al. 2004; Kristensen and Kijl 2010), or to *legitimize* the investment costs (Desmarais et al. 1997).

	Understanding	Control	Development	Legitimation
Purpose	Li and Gupta 2009; Lin and Lee 2006; Majchrzak et al. 2006; Schubert and Williams 2011	Markus and Tanis 2000; Muller et al. 2009; Rosemann and Wiese 1999	Davenport et al. 2004; Kristensen and Kijl 2010; Seddon et al. 2010	Desmarais et al. 1997; Shang and Seddon 2002; Tallon et al. 2000

The distinction of different **levels of analysis** is widely used in related works and can be found in many multidimensional approaches in order to get a complete picture of the IS impact (e.g. Gable et al 2003; Grover et al. 1996). The differentiation of *individual* and *organization* is mostly common. Furthermore, these characteristics are also considered separately (individual: e.g. Davis 1989; Venkatesh et al. 2003, organization: Kettles and David 2008; Seddon et al. 2010). An additional view across multiple instances is the *group* view (e.g. Duivenvoorde et al. 2009; Fitzgerald 1998; Ifinedo 2006; Janhonen and Johanson 2011; Joosten et al. 2011; Kurian et al. 2000).

	Individual	Group	Organization
Level of analysis	Al-adaileh 2009; Davis 1989; DeLone and McLean 1992; Gable et al. 2003; Gable et al. 2008; Grover et al. 1996; Hsu and Lin 2008; Venkatesh et al. 2003	Duivenvoorde et al. 2009; Easley et al. 2003; Fitzgerald 1998; Ifinedo 2006; Janhonen and Johanson 2011; Joosten et al. 2011; Kurian et al. 2000	DeLone and McLean 1992; Gable et al. 2003; Gable et al. 2008; Grover et al. 1996; Kettles and David 2008; Shang and Seddon 2002; Seddon et al. 2010

The fourth dimension – the **perspective** of the entities – comprises different target groups within the company that dominate the interests of an evaluation. First, *management* is interested in the incurred costs as well as the realized benefits, in order to be able to justify investments (e.g. Majumdar et al. 2013; Smits and Mogos 2013). For the *user*, however, it is important to understand the benefits of using a system for his or her daily work (e.g. Tarafdar et al. 2010; Al-adaileh 2009) and to identify opportunities for further development (e.g. Majchrzak et al. 2006). Due to the social functionalities of an ESS, the users are more and more interested in usage statistics (how many likes, views, visits does my page have?). The *platform owners* who are responsible for a system are interested in the system's performances and processes as well as the generated business value (e.g. Richter et al. 2013).

	User	Platform owner	Management
Perspective	Al-adaileh 2009; Majchrzak et al. 2006; Tarafdar et al. 2010; Venkatesh et al. 2003	Davenport et al. 2004; Richter et al. 2013; Rosemann and Wiese 1999; Seddon et al. 1999	Kristensen and Kijl 2010; Majumdar et al. 2013; Shang and Seddon 2002; Smits and Mogos 2013; Tallon et al. 2000

The role of system usage and its relationship to its success is considered controversial. Some authors take the position that usage itself means success or leads to success and thus that user acceptance and usage are important parts of the system evaluation (e.g. Davis 1989; Du and Wagner 2006). On the other hand, some researchers see no need to consider use and user satisfaction within evaluation. They either see them merely as a predecessor and result of the IS's impact (e.g. Gable et al 2008; Rai et al. 2002), or

recommend concentrating on the business value (e.g. Melville et al. 2004). Although the exact relationships are controversial, the **evaluation objects** *usage* (measures are e.g. usage statistics, usage behavior, degree of cross-linking, user requirements) and *business value* (e.g. productivity, cost or time savings, ROI, cost-benefit-ratios, process impacts) are closely linked; they are often considered integral (e.g. DeLone and McLean 2002; Soh and Markus 1995). Usage analysis appears mainly in the ESS literature, as with ESS, a variety of user-generated content is created (e.g. Hsu and Lin 2008; Kane et al. 2014; Trimi and Galanxhi-Janaqi 2008).

	Usage	Business value
Evaluation objective	Davis 1989; DeLone and McLean 2003; Du and Wagner 2006; Raeth et al. 2012; Richter 2013; Soh and Markus 1994; Trimi and Galanxhi-Janaqi 2008; Venkatesh et al. 2003	DeLone and McLean 2003; Gable et al. 2008; Melville et al. 2004; Richter et al. 2013; Soh and Markus 1995; Stefanou 2001; Tallon et al. 2000

The underlying **data** can roughly be divided into *qualitative* (e.g. Kügler et al. 2013; Weill 1992) and quantitative. Additionally, a distinction of the quantitative data into *monetary* (e.g. Desmarais et al. 1997) and *nonmonetary* is often used in related studies (e.g. Muller et al. 2009; Raeth et al. 2012). Qualitative data are often used for the purposes of understanding and development. As quantitative data are easier to compare, they are well suited for control and legitimacy. For legitimation, monetary data are most commonly used. In a subsequent analysis, the collected qualitative data can also be quantified in order to allow for comparison. However, mostly different types of data are combined for achieving meaningful evaluation results (e.g. DeLone and McLean 1992; Gable et al. 2003). In regard to the evaluation object, an accumulation of analysis of non-monetary-quantitative usage statistics can be seen in ESS publications. This can be traced back to the possibility to simply collect data at low cost (Behrendt et al 2014).

	Qualitative	Quantitative nonmonetary	Quantitative monetary
Data	DeLone and McLean 1992; Gable et al. 2003; Janhonen and Johanson 2011; Kügler et al. 2013; Majchrzak et al. 2006; Seddon et al. 1999; Stefanou 2001; Weill 1992	DeLone and McLean 1992; Gable et al. 2003; Janhonen and Johanson 2011; Muller et al. 2009; Raeth et al. 2012; Rosemann and Wiese 1999; Seddon et al. 1999; Stefanou 2001	DeLone and McLean 1992; Desmarais et al. 1997; Gable et al. 2003; Rosemann and Wiese 1999; Seddon et al. 1999; Stefanou 2001

Closely connected to the data dimension, the **data gathering** methods are also relevant. Two of the most applied methods are *interviews* (e.g. Ashurst et al. 2008; Shang and Seddon 2002) and *surveys* (e.g. Majchrzak et al. 2006; Urbach et al. 2009b). Furthermore, *log file extractions* are used for the simple collection of usage data (e.g. Easley et al. 2003; Muller et al. 2009), which is widely used in the ESS literature. The *extraction of content* is another method to obtain an understanding of a system's use and processes (e.g. Schubert and Williams 2011). Furthermore, data can be gained by means of *process and usage observations* (e.g. Kügler and Smolnik 2013; Richter et al. 2013; Weill 1992). In addition, *financial figures can be calculated*; these must be procured from the accounting department or other instances (e.g. Desmarais et al. 1997).

	Interviews	Surveys	Log files extraction	Content extraction	Process and usage observation	Financial calculation
Data gathering	Ashurst et al. 2008; Shang and Seddon 2002; Steinhüser et al. 2011	Majchrzak et al. 2006; Janhonen and Johanson 2011; Urbach et al. 2009b	Easley et al. 2003; Janhonen and Johanson 2011; Muller et al. 2009; Steinhüser et al. 2011	Richter and Riemer 2013b; Schubert and Williams 2011	Kügler and Smolnik 2013; Richter et al. 2013; Weill 1992	Desmarais et al. 1997; Rosemann and Wiese 1999; Stefanou 2001

The **context** addresses the data´s relationship; that is, what the results should be compared to. Grover et al. (1996) and Shang and Seddon (2002) call this dimension referent and define three different characteristics. Basically, the data can be compared with an old or similar system's data in the same company, but also with comparable systems in similar companies (*comparative*) (e.g. Pirkkalainen et al. 2009; Smits and Mogos, 2013). Otherwise, the interpretation of the results compared to a theoretical ideal

system or best practice is another variant (*normative*) (e.g. Irani and Love 2001; Markus and Tanis 2000). The third possibility is that the system is evaluated in the course of its own development over time (*improvement*) (e.g. Janhonen and Johanson, 2011; Melville et al. 2004). Also the characteristics of this dimension is equal in both ESS and IS.

	Comparative	Normative	Improvement
Context	Grover et al. 1996; Muller et al. 2009; Pirkkalainen et al. 2009; Smits and Mogos, 2013	Grover et al. 1996; Irani and Love 2001; Markus and Tanis 2000; Shang and Seddon 2002	Grover et al. 1996; Janhonen and Johanson, 2011; Melville et al. 2004; Shang and Seddon 2002

Constructs and Principles of Form and Function

A Conceptual Framework of ESS Evaluation

The following section presents the abstract 'blueprint' (Gregor and Jones 2007), with its construct entities and principles of form and function. The framework's compilation is based on the theoretical foundations, as established in the justificatory knowledge section. Table 1 shows the conceptual framework, with eight dimensions and the corresponding characteristics.

Table 1. A Conceptual Framework for ESS Evaluation

		Characteristics							
Dimensions	**Time** *When to evaluate?*	Ex ante		Ongoing		Ex post			
	Purpose *Why to evaluate?*	Understanding		Control		Development		Legitimation	
	Level of analysis *Which level to evaluate?*	Individual		Group		Organization			
	Perspective *Who to evaluate for?*	User		Platform owner		Management			
	Evaluation object *What to evaluate?*	Usage			Business value				
	Data *What kind of data to use?*	Qualitative		Quantitative nonmonetary		Quantitative monetary			
	Data gathering *How to get the data?*	Interviews	Surveys	Log files extraction	Content extraction	Process and usage observation	Financial calculation		
	Context *In which relationships are the data?*	Comparative		Normative		Improvement			

The construct of the framework consists of eight *dimensions* that are relevant to ESS evaluation. Each dimension has several *characteristics*. For an evaluation each dimension and at least one or more characteristics are to be considered. It can be seen that all dimensions and characteristics occur in both the ESS and IS evaluation literature, however, the focus on the characteristics differs. For instance, the characteristic "usage" of the dimension "evaluation objective" is frequently applied in ESS evaluations, which can be traced back to the amount of user generated content and the possibility to simply extract log files at low costs. In the section Principles of Implementations and Expository Instantiation typical combinations of the characteristics are summarized in ESS *evaluation classes*. Dimension, which defines different proportions in terms of the overall system, is also used in other models (DeLone and McLean 1992; Gable et al. 2003) to describe a system's different effects. In contrast, we refer to dimensions as integral parts of the evaluation – similarly to Grover et al. (1996).

Being aware of the dimensions and their particular characteristics is essential for the ESS evaluation's design and implementation. The conceptual framework provides a template with supplementary explanations and principles for the evaluation design. Every dimension has a question that should be considered in order to achieve an effective and efficient ESS evaluation: *When to evaluate?* (time), *Why evaluate?* (purpose), *Which level to evaluate?* (level of analysis), *Who to evaluate for?* (perspective), *What to evaluate?* (evaluation objective), *What kind of data to use?* (data), *How to get the data?* (data gathering) and *In which relationships are the data?* (context). Each dimension and characteristic has a different influence on the evaluation design.

Validation with an Interview Study

Approach

To investigate the ESS evaluation applied in practice, we adopted a qualitative approach (Schultze and Avital 2011). 31 Semi-structured interviews from 29 companies were conducted between December 2011 and April 2014. This allowed us to make a confirmatory assessment of the dimensions and their particular characteristics (Spencer et al. 2003). It also enabled us to identify relationships between them, as well as to understand the general approach of an ESS evaluation. All participating companies have broad ESS application experience. The participants have been approached at conferences, exhibitions, in web communities or by personal recommendations. The companies' employees numbered between 10 and more than 400,000. The spectrum of industries ranges from automotive to banks, communication services, insurance, (business and media) consulting, military, energy, healthcare, and chemicals. Prior to the interviews, we developed an interview guide to support the conversations with interviewees (Bryman and Bell 2007). The interview guide contained 32 questions in four different categories, including questions about the person and the company, about ESS experience, about measuring ESS success, and about usage's influence on success. This enabled us to get an idea of a participant's experience and, simultaneously, to detect different contexts of statements. The interviews were recorded and had an average length of 45 minutes. We transcribed the interviews and subsequently coded the text documents. Thereby we categorized the different codes under the respective question. Additionally to classifying the codes into the different dimensions and characteristics, we also counted their frequency. Finally, we allocated the different characteristics to evaluations classes (see "Principles of Implementation and Expository Instantiation"). Although 31 interviews do not allow us to make a representative statement, they are a valuable indicator of the weighting by which we could obtain information about their importance and the focus of the ESS evaluation. Because the interviews have been conducted in German, the quotations in this paper were translated into English (Regmi et al. 2010).

Results

In 26 out of 29 companies, the ESS has been evaluated at least once. The need for an evaluation was high and the benefit was recognized (Quote of interviewee 9 further i09): "If it is important that the solution [ESS] in the company delivers an added value, then this added value needs to be evaluated. When you see that the added value is too low, you have to take action in order to increase this. And in turn this must be evaluated again."). All eight dimensions and their characteristics identified in the literature were confirmed in practice. Each characteristics of a dimension has an influence on the evaluation design and implementation. Thus, in the following, the dimensions and characteristics are explained and illustrated in their influence. This allows for a better understanding and description of the principles of form and function. The interviews also allowed us to identify and examine typical instantiations in form of six main evaluation classes, which can be found later in the principles of implementation and expository instantiation.

Time and purpose. In many interviews, we found a close relationship between the dimensions time and purpose. Additionally, it can be seen that time and purpose determine whether the evaluation is pre-formative (examine the conditions), formative (constructively, creatively), or summative (collectively, summarizing). Ex ante analyses are generally formative, while an ex post analysis is summative owing to a lack of design aspects. However, ongoing analysis can be both formative and summative. The purposes themselves are also not independent. Thus, for instance, control, development, or legitimation always

develop an understanding and vice versa. However, because defining a primary purpose determines the evaluation design, the distinction is necessary (Stockman 2011).

In one company, ex ante evaluation results were used to obtain the required budget for an official ESS project from management (i30: "Before the introduction, we had calculated how many license costs arise with Confluence, which hardware and server we need, and how much the staff effort in hours would be for the introduction. We then presented this to management."). But the prerequisites for the implementation of an evaluation depend on organizational conditions. For instance, in some cases, the ESS was introduced unofficially, bypassing any ex ante legitimation (i06: "It was never a decision from above or from a designated authority that said: We use it now with support from management, and in a year, we see how far we have come. This was not the case. It was a completely bottom-up approach."). Furthermore, the basic system design requirements were clarified ex ante. In one example, this was done on the basis of a pilot system from external evaluators (i27: "We have thoroughly evaluated the pilot of the ESS [...] just to show the platform responsible the use and benefits. Then they decided to readjust some points and subsequently to roll out the system.").

Ongoing and ex post evaluations were more common than an ex ante evaluations. If the ESS implementation is at an early stage, the interviewees sought to achieve an understanding of the usage by qualitative data firstly (i09: "An evaluation of the wiki took place. [...] Maybe not exactly with the focus to measure the success, but to get a snapshot of how the wiki has been integrated into the work processes, or how the wiki is perceived by the employees."). It can be seen that especially the platform owners wants to understand how ESS is adopted and in which way it used by the employees. Furthermore, ongoing studies were used for continuous system development. Similarly to the purpose of understanding, mainly qualitative data were supplemented with usage data (i18: "In the course of a weak point analysis, we combined usage tracking and user interviews. We have tried to correlate them in order to identify weaknesses and requirements, and to set priorities for the development of the new release."). At the same time, ongoing evaluation is used for the control of the fulfillment of stated targets (i01: "Our goal was that at least one-third of the users participate as authors. This was achieved quickly and we were able to measure that too. Meanwhile, 60% of all users are authors.").

Target achievement control was also ex post – many companies face issues in measuring the financial effects (i11: "Our success measurement would always be ex post. Actually, we have only measured the usage ex post, because ROI is very difficult in this field. [...] and this I can only measure after a certain time, not during the project implementation phase."). Despite the difficulty of proving direct effects, some participants were able to implement the ex post justification. Among other things, this was done by calculating costs savings such as travel or hardware costs. The justification to management was in some cases supplemented with additional qualitative data (i15: "Last year, I queried the business-relevant facts and presented the results to management. Here, I compared the cost of the platform with the business benefit to the company. If the business benefits are higher than the costs, then the company has a positive impact.").

Level of analysis. All levels of analysis (individual, group, and organization) were also found in the interviews. Here, for instance, data such as usage statistics from users are mostly summarized to groups at certain granularities (i04: "[...] there are cases in which the individuals operate in a way that is unique to a department. [...] Someone works on a project and edits, searches, or works in this context. Since you can just abstract from the individual level and raise the whole thing to a group."), or simply their entirety for the organization is evaluated. But data policies were often reviewed in advance for the data gathering on individual users, as especially for social software, it is critical to protect personal data (i01: "The data collection refers to the whole amount of users. The data analysis cannot be implemented as personalized at the employee level owing to work council agreements."). However, the analysis can be implemented with anonymous data (i26: "[...] But it is done anonymously and I can already evaluate this anonymous data, which are not about the individual employees.").

Perspective. The three distinct perspectives of the evaluation could also be confirmed. First, in many cases, the investment had to be justified in order to explain the invested time and costs, or to get a new budget (i21: "I have to justify it to my colleague in management, who has nothing to do with technology; she comes from the creative side."). However, it is important for users to understand the benefits of using ESS in their daily work and to identify further development opportunities (i26: "We do regular surveys: How is the user experience? [...] We are very, very interested in the results, because of course they flow

back into product development."). The platform owners themselves are standing in between and want to know both the usage and the business value. So the perspective has an influence on the evaluation design and workup and presentation of the evaluation results. (i04: "First of all, we [platform managers] are interested in evaluating the success, because we want to run the project well and more successfully. Sometimes we present the results to management or our work colleagues. But often there are also users who want to determine, for instance, where it's best to host their information – either in the wiki or maybe in SharePoint – and how many views there are.").

Evaluation object, purpose and data. The subdivision of the evaluation object into usage and business value, which had emerged from the literature, could not only be confirmed but plays a special role in ESS evaluation practice. Owing to an abundance of user-generated content and the high degree of employee networking, the evaluation of usage forms part of many evaluations. Almost all the participating companies collect usage statistics (22 out of 29) and analyze usage (see Appendix A). This is promoted among others by the simple and inexpensive data gathering opportunities that analysis tools present. Nearly every ESS provides analysis tools with which it is possible to track the activity of users in the form of traffic, page views, number of users, logins, or blog posts (i11: "We have some automated control mechanisms based on database queries in the tool."). However, the impact and the significance of usage in terms of the business benefits are difficult to prove (i01: "What is meant by success in business is usually money. Therefore, the correct measurement of success collects the monetary aspects, but this is in my opinion difficult in terms of enterprise 2.0. We must therefore use other methods."). Otherwise, the evaluation of usage is of great importance, primarily for the purposes of development and understanding. For the purpose of legitimation, an evaluation of business value is most often used. In this context, it may be that the interests in terms of the evaluation objective change also over time (i25: "In the next phase, we first really want to know if there is activity on [the platform]. If there is regular activity on it, you suddenly become interested in how it can deliver economic benefits. [...] interests also move the more mature a system is."). A combination of usage and business value is also often used to emphasize the benefits.

Data, data gathering, and purpose. The characteristic of the evaluation object decides about the kind of data used in the evaluation. The special focus on usage in ESS evaluation is also reflected in the measures and data gathering methods (see Appendix A). The selection of the required data and measures also indirectly determine the data gathering methods. However, there may be several options for one measure. Thus, for instance, monetary data can be also collected by means of estimates in interviews or surveys. There is also the question whether surveys or interviews are better. This depends for instance on whether much comparable data are needed (surveys) or whether deeper expert knowledge (interviews) is important. Purpose is again an important influence dimension. For the understanding or development approaches, qualitative information tended to be used (i25: "It depends on the target horizon [...]. It is very difficult to use quantitative methods in a relatively new research field. More qualitative evaluations are suitable for this."), while for the legitimation and control, meaningful quantitative data are more likely (i31: "Yes, of course it costs money, so ROI is also a big topic. [...] And management prefers a quantifiable sum, even if not all positive effects can be monetized."). Again, there may be mixed forms or combinations for meaningful evaluation results.

Context and data. Finally, the question is how to interpret the data. For instance, a utilization of 60% can merely be taken into account when this is compared to other data. If an old tool was used only 30%, the 60% can be rated (independently of other interpretations) positively. Basically, the results can be compared with those of old systems in the same company, but also to comparable systems in similar companies (*comparative*). Unfortunately, comparative data are often missing, which could prevent a meaningful assignment of an enhancement compared to the old system (i23: "Unfortunately, I have no comparative data to which I can compare the new and old intranets. The data should have been collected shortly before the conversion to see what has changed, but my predecessor did not do this."). In one case, surveys were used after each conversion from one system to another in order to compare them. The interpretation of the results compared to a theoretical ideal system or best practice is another variant. But this normative evaluation was applied only in one case where an external evaluator conducted surveys in several companies; he could therefore prepare a benchmark and best practice comparison. A system is often evaluated in the course of its development over time (*improvement*). However, the significance of an evaluation is higher when the results can be compared to results of other systems.

Testable Propositions

Testable propositions seek to draw conclusions about a design theory's utility (Venable 2006) and thus to support its validity (Offermann et al. 2010). It is essential that the kernel theories can be put to practical use (Walls et al. 1992). Through our interviews, we were able to meet this demand and to validate the conceptual framework. At the same time, this comparison of literature and practice enabled us to make testable propositions (Gregor and Jones 2007).

When analyzing the interviews, it became clear that the dimensions and their particular characteristics appear in any ESS evaluation. Some interviewees consciously considered this, while others applied an unsystematic evaluation approach without an a priori design. We found that consciously examining the different dimensions and characteristics result in more efficient and effective evaluations. At this point, our design theory can provide valuable support. Concerning utility (Venable 2006) and validity (Offermann et al. 2010), and based on the kernel theories as well as the interview statements, we make the following testable propositions:

1. Consciously considering all dimensions and characteristics results in a complete and consistent ESS evaluation. The effectiveness and thus the quality of an evaluation's results can be increased.
2. Consciously considering all dimensions and characteristics results in a resource-saving ESS evaluation. The efficiency and thus the productivity of an evaluation's design and application can be increased.

Principles of Implementation and Expository Instantiation

The principles of implementation – also described as a design method (Walls et al. 1992) – are of particular relevance for design theories. In our case, we indicate how to apply the ESS evaluation framework, which shows the ESS evaluation dimensions and characteristics. Being aware of these dimensions and characteristics results in a better understanding of the evaluation, which, in turn leads to a more efficient evaluation design process. Depending on the specification of different dimensions, the process of designing an evaluation based on our design theory differs in each organization's situation. Some particular characteristics are predefined by given circumstances, while the remaining ones are assessed in their interrelationships. For instance, an evaluation might be requested by management in order to show the effects of ESS on the project execution. This means that the perspective (*management*) and the purpose (*understanding*) are given. Furthermore, the dimension of time (*ongoing*) is usually given by the current project phase. However, the context must also be considered. An ex post view of the current system can also be an ex ante view for a future system. In our example, time, purpose, and perspective are given. Once the exact requirements are clarified, one must consider which characteristics of the remaining dimensions apply and what this means for the evaluation design. For instance, which level of analysis is most appropriate for the purpose? Is usage or business value more important, or a combination of both? Which data are needed for this and how to best collect them? What kind of comparison is appropriate? Which comparative data are available?

The identified dimensions can be divided into two groups, which are organizational (time, purpose, level of analysis, and perspective) and data-related (evaluation object, context, data, and data gathering). There is a very strong dependence between the data-related and the organizational dimensions. The characteristics of the organizational dimensions mainly affect the choice of data-related characteristics. For the evaluation design, this means that it is better to first clarify organizational dimensions and then the data-related dimensions. The developed framework can therefore guide the evaluation design and provide useful orientation. To illustrate the principles of implementation, we have studied typical combinations of characteristics in the framework gained from our empirical interview data. Our goal was to aggregate the characteristics under consideration of specific situations (Offermann et al. 2010). We were able to identify six primary evaluation classes (see Figure 1). An overview of each class's measures, which were gained from the interviews, can be found in Appendix A.

The 26 out of 29 companies participating in our study conducted 63 ESS evaluations. Accordingly, each company has implemented 2.42 different ESS evaluations on average. We assigned each single evaluation approach to one class. The assignment revealed for example that 21 companies evaluate their ESS in order to gain an understanding of the usage and effects (Class III). This can be attributed to the way how an ESS

is adopted by its users which is not a priori defined and will only be shown in the course of time. The high number of evaluations assigned to Class III also allows conclusions about the ESS maturity in the participating companies. Once an understanding has been reached, system improvements and a review of the company's goals can be addressed. Another finding to be highlighted is that only two companies justified their investment (Class VI) by monetary data. Instead, difficulties in clearly assigning monetary effects to an ESS, leads companies to consult qualitative data and usage statistics when it comes to justify their investment. The relatively low amount of ex ante evaluations can be traced back to the bottom up diffusion of ESS which was also found in other studies (Richter and Riemer 2013a), whereas ex post evaluations occur equally rarely due to the fact that the ESS were just recently introduced in many companies.

*n=63 [26 companies]**	**Class I: Investment estimation** *n=6*	**Class II: Requirement analysis** *n=5*	**Class III: Understanding usage and effects** *n=21*	**Class IV: Improvement of system or process** *n=11*	**Class V: Controlling achievement of targets** *n=13*	**Class VI: Investment justification** *n=7*
Time	Ex ante *(n=6)*	Ex ante *(n=5)*	Ongoing *(n=21)*	Ongoing *(n=11)*	Ongoing *(n=11)*; Ex post *(n=3)*	Ongoing *(n=4)*; Ex post *(n=3)*
Purpose	Legitimation *(n=6)*	Development *(n=5)*	Understanding *(n=21)*	Development *(n=11)*	Control *(n=13)*	Legitimation *(n=7)*
Level of analysis	Organizational *(n=6)*	Group *(n=4)*; Individual *(n=3)*; Organizational *(n=1)*	Organizational *(n=13)*; Group *(n=10)*; Individual *(n=8)*	Organizational *(n=8)*; Group *(n=5)*; Individual *(n=4)*	Organizational *(n=10)*; Group *(n=5)*; Individual *(n=2)*	Organizational *(n=4)*; Group *(n=3)*
Perspective	Management *(n=5)*; Platform owner *(n=2)*	Platform owner *(n=5)*	Platform Owner *(n=21)*; Management *(n=5)*; User *(n=2)*	Platform owner *(n=11)*; User *(n=1)*	Platform owners *(n=11)*; Management *(n=6)*	Management *(n=7)*
Evaluation objective	Business value *(n=6)*	Usage *(n=5)*	Usage *(n=21)*; Business value *(n=13)*	Usage *(n=11)*; Business value *(n=3)*	Usage *(n=10)*; Business value *(n=7)*	Usage *(n=5)*; Business value *(n=4)*
Data	Quantitative monetary *(n=6)*; Qualitative *(n=4)*	Qualitative *(n=5)*	Qualitative *(n=18)*; Quantitative nonmonetary *(n=18)*; Quantitative monetary *(n=8)*	Qualitative *(n=11)*; Quantitative nonmonetary *(n=7)*; Quantitative monetary *(n=2)*	Quantitative nonmonetary *(n=11)*; Quantitative monetary *(n=7)*; Qualitative *(n=5)*	Qualitative *(n=6)*; Quantitative nonmonetary *(n=6)*; Quantitative monetary *(n=2)*
Data gathering	Financial calculations *(n=6)*; Interviews *(n=3)*; Surveys *(n=2)*	Interviews *(n=4)*; Surveys *(n=1)*	Log files extraction *(n=17)*; Interviews *(n=10)*; Surveys *(n=9)*; Financial calculations *(n=5)*; Process and usage observation *(n=4)*; Content extraction *(n=3)*	Surveys *(n=7)*; Interviews *(n=6)*; Log files extraction *(n=5)*; Process and usage observation *(n=2)*; Financial calculation *(n=2)*	Log files extraction *(n=12)*; Financial calculation *(n=4)*; Surveys *(n=4)*; Interviews *(n=3)*; process and usage observation *(n=2)*	Log files extraction *(n=6)*; Interviews *(n=4)*; Surveys *(n=4)*; Financial calculations *(n=1)*
Context	Comparative *(n=6)*; normative *(n=2)*	Comparative *(n=5)*; Normative *(n=3)*	Improvement *(n=20)*; Comparative *(n=6)*; Normative *(n=3)*	Improvement *(n=10)*; Comparative *(n=2)*; Normative *(n=1)*	Improvement *(n=11)*; Comparative *(n=5)*	Improvement *(n=6)*; Comparative *(n=2)*; Normative *(n=1)*

** 63 ESS evaluations from 26 out of 29 companies*

Figure 1. Classes of ESS Evaluation

Because the classes were derived from the interviews, they reflect real instantiations in practice (Dubin 1978). While March and Smith (1995) as well as Hevner et al. (2004) see an instantiation as an artifact of design science, Gregor and Jones (2007) argue that it refers to a specific situation in which an instantiation is an exemplary declaration but does not represent an artifact itself. Therefore, while it is an important ingredient, it relates only to specific situations. The classes show specific application descriptions, so that the evaluation design is illustrated in particular situations (Offermann et al. 2010). However, to verify a conscious application in practice, the conceptual framework and the classes were presented in a small-sized, a mid-sized and a large company. We sought to represent and present our design theory (Gregor and Jones 2007) to participants who would apply it in their ESS evaluation. At the outset, without showing the framework, we asked participants about possible dimensions and characteristics. Here, only an empty dimension framework has been shown to verify to what extents the characteristics are consciously involved during an evaluation. In all three cases, only some characteristics were named, whereby it was shown that they are not fully taken into account as expected. Even the dimensions were only partially involved in previous evaluations by all participants.

After the complete framework was shown and explained, the participants confirmed the characteristics. All three participants indicated that the framework provides an overview of the evaluation scope for its design and creates an awareness of the different influences. However, accurate application scenarios were said to be lacking. These were then covered by the subsequent presentation of the six classes. Thereby, the classes have demonstrated the potential and the possibilities for further evaluations, which have not yet been carried out (participant 1: “Through the classes, I have an orientation of how I can address the evaluation and which steps I need to take. But they also show which evaluations should be used for our system.”). The characteristics in the evaluation design then need to be considered again in detail (participant 2: “The classes provide a useful guide, but then you have to individually look at the various characteristics and dimensions to see if these fit the particular situation.”). In addition, the Framework and the classes were presented to two research groups in order to finally validate the consistency (Walsham 2006). In summary, we have critically discussed instantiations in practice in the form of five presentations. It was shown that the conceptual framework and the six evaluation classes raise awareness of the different influencing factors and provide guidance in evaluation design.

Artifact Mutability

According to Gregor and Jones (2007), ensuring mutability in design theories is challenging, considering the dynamism of the technology sector and ever-shorter innovation cycles. This mutability concerns two kinds of changes. On the one hand, the artifact must be able to adapt to different application scenarios. On the other hand, being able to flexibly adapt the theory itself in its form and shape is also important (Offermann et al. 2010). For a design theory of ESS evaluation, this means: (1) to what extent are the dimensions and characteristics of the framework artifact customizable, and (2) to what extent can the class compilation be flexibly extended and/or adjusted over time.

When creating our design theory, we anticipated the possibility of adapting it to specific situations a priori. As a result, the framework has been designed for a variety of combinations and evaluation approaches. However, this implies that further changes are possible, should the situation or circumstances require it. In specific dimensions, a more granular view of the characteristics is conceivable. For instance, several types of groups could be distinguished in the level of analysis, or the dimension data could be further subdivided. Since the characteristics must always be interpreted and adapted, as required by the application situation, a more granular view does not provide added value for the general framework, but restricts its clarity. The six presented evaluation classes represent typical application scenarios where subclasses and situational variations in characteristics are possible. The classes’ modular generic structure thus allows one to adjust and supplement them at any time. However, possible impacts on the remaining dimensions and characteristics must be considered.

Over time, the interest with respect to the ESS change and, with that, the alignment of the evaluation design accordingly. After a new system has been introduced, it is important to reach system usage critical mass. The evaluation should then focus on usage types. Once the ESS is used successfully, the interest moves from the usage to the business value. Similar to other models (Markus and Tanis 2000; Tallon et al. 2000), the evaluation classes are designed for different project and lifecycle phases. Experiences from earlier evaluations can be taken up and considered for future approaches. These may include changes in

the data gathering (were surveys the right choice, or do I need interviews next time?), in the data (were the right questions asked? Do they need to be supplemented?), or the level of analysis (is looking at the user level appropriate, or should I expand my evaluation on the entire organization?). Furthermore, comparative data from previous states or other systems can be considered. Mutability over time is thus given and can contribute to optimizing the evaluation.

Conclusion

ESS evaluation is a complex process owing to a plethora of different interests and perspectives. There are several stakeholders who have different interests in an evaluation at different times. At the same time, very few studies provide an overview of factors to be considered or provide guidelines on how ESS can be evaluated comprehensively. We have argued that, despite the usefulness of existing studies in the IS field which provide such guidelines (Grover et al. 1996, Seddon et al. 1999), the emergence of ESS makes it necessary to re-examine and possibly adapt or revise them. Therefore, our study shed light on the various design options for ESS evaluations making it a valuable contribution for a company's ESS management. Based on a comprehensive literature analysis, as well as an interview study, we suggest a conceptual framework intended as decision support during the design phase of ESS evaluation. While the framework builds the orientation and starting point, it is accompanied by six evaluation classes that represent typical application instantiations and can be understood as implementation principles. Owing to the need for a theoretical foundation and the coexisting challenge of practicality (Rosemann and Vessey 2009), we applied the research framework for the creation of a design theory of Gregor and Jones (2007). Table 2 provides a summary of our design theory, classified according to Gregor and Jones's (2007) components.

Table 2. Summary of the Design Theory for the Evaluation of ESS

Type	Component
Purpose and scope	The artifact of the design theory is an evaluation model, provided for both practitioners and researchers to individually design and implement an evaluation approach for ESS. This individuality is enabled by the inclusion of various views and influence factors for different situations and interests.
Justificatory knowledge	51 publications from ESS as well as other types of IS evaluation were taken into account, to identify and analyze the existing evaluation dimensions and characteristics.
Constructs	Dimensions and dimension groups, characteristics, and classes of implementation.
Principles of form and function	A conceptual framework with eight dimensions that includes different questions for an ESS evaluation and related characteristics. Each of these dimensions and characteristics should be considered, as to whether they have a lesser or greater influence on the evaluation.
Testable propositions	Comparing the literature to practice allows for the derivation of two testable propositions. In summary, the utility of the design theory is caused by increases in evaluation effectiveness and efficiency.
Principles of implementation	The approach to implementation differs according to the specific situation in the company. Generally, the organizational dimensions should be clarified before the data-related dimensions. Six main evaluation classes are described to represent typical approaches.
Expository instantiation	The conceptual framework and the classes were presented and discussed in five cases with participants from science and practice. The six evaluation classes represent instantiations in specific situations.
Artifact mutability	The design theory is flexibly designed from the ground up for specific situations and over time. A supplementation of dimensions and characteristics is possible; in this case, the effects on other dimensions must be weighed. For the combination of the characteristics, other subclasses besides the six main classes are conceivable.

Our results show that different dimensions should be considered when evaluating ESS. The conscious consideration of these dimensions allows for a more effective and efficient evaluation design. The identification and composition of the characteristics in six evaluation classes shows a change in different project and lifecycle phases of the ESS – not only the interests of an evaluation shift, but the overall design and composition of the characteristics of the dimensions. Hence, evaluations are subject to a dynamic process and must always be redesigned over time.

It can be seen that, all dimensions and characteristics occur in both the ESS and IS evaluation literature. However, the focus on the characteristics differs which was also confirmed by the results of the interview study and is displayed in the evaluation classes. Thus, we were able to show the peculiarities of an ESS evaluation. For instance, *understanding* ESS usage and effects (cf. evaluation class III) seems to be essential for ESS experts. The malleability of ESS implies that its use scenarios will only be explored and established over time. Users have to make sense of and appropriate the platforms in their daily practices. Furthermore, *usage* statistics are the most frequently gathered data. We explain this by the large amount of user-generated content and the simplicity of collecting usage data via analysis tools at low costs. These circumstances make the measuring very easy, especially for platform owners who want to learn about the adoption rates of ESS. However, the meaningfulness of usage statistics with respect to the business value requires additional evaluations. Additionally it can be seen that, compared to other types of IS, ESS is often diffused bottom up in the company whereby a comprehensive ex ante evaluation is skipped.

In summary, we have shown that an evaluation is an individual and situation-dependent activity. Its diversity is reflected in the different models in the literature and several application scenarios in practice. With few exceptions, the existing models do not sufficiently address all the views on an evaluation, but are limited to few dimensions and isolated characteristics. Here, our framework unites these different perspectives and provides a novel approach in form of a design theory. Although we do not claim that our study is representative, its explorative character nevertheless produced interesting results. Thus, in order to better validate the proposed framework, we are in the process of collecting more datasets from other cases. Even if the research field is ESS, we assume that cognitions can be transferred to other types of IS which we want to investigate in further research.

Appendix A. Measures for the ESS Evaluation

Table 3. Measures for the ESS Evaluation

Data gathering	Measures
Financial calculations	Travel cost savings; IT cost savings; costs per user; hardware costs, software license costs; costs of employee effort
Content extraction	Text length; number of links; number of images; scenarios of use; number of contact requests, direct messages, and comments between users; sentiments
Log files extraction	Number of: blog posts, community spaces, authors, attachments, visits, edits, log-ins, user, new users, messages per day, unique visitors, comments, blog followers, average comments on a blog or discussion forum, unique users and hits per time period, sessions, wiki pages created per day, posts, readers of a post, praises per post; Development of use; posts with most readers; average time per user per visit; session time; top 10 user groups (regarding number of articles, members, page views); Email frequency; number of emails; correlation of email frequency and blog posts from one person
Interviews	User requirements for the platform; usage behavior; use case validation / user satisfaction with the processes or tools; review of the tool; use cases ROI of use cases; opportunity proceeds of projects; ROI; cost-benefit ratio
Surveys	User satisfaction with the platform; usage types of the tool; frequency of use; applicability of the tools; users' knowledge of the tools' possibilities; satisfaction with availability through the new tools; self-assessment of the affinity in using

	new tools; usability benefit; usage behavior; use case validation; Effort of working with the tools; individual business value; ROI; money saved; generated new revenue; new customer acquisition; perceived benefits for the organization; business value
Process and usage observations	Adjusted ideas; intensity of collaboration; degree of cross-linking; genres of use Time spent reading, writing and answering; number of implemented ideas; number of awards

References

Al-adaileh, R. M. 2009. "An Evaluation of Information Systems Success: A User Perspective - the Case of Jordan Telecom Group," *European Journal of Scientific Research* (37:2), pp. 226-239.

Aral, S., Dellarocas, C., and Godes, D. 2013. "Introduction to the Special Issue - Social Media and Business Transformation: A Framework for Research," *Information Systems Research* (24:1) pp. 3–13.

Ashurst, C., Doherty, N. F., Peppard, J. 2008. "Improving the Impact of IT Development Projects: the Benefits Realization Capability Model," *European Journal of Information Systems* (17:4), pp. 352-370.

Baskerville R., and Pries-Heje, J. 2010. "Explanatory Design Theory," *Business & Information Systems Engineering (52:5)*, pp. 259-271.

Behrendt, S., Richter, A., and Trier, M. 2014. "Mixed methods analysis of enterprise social networks," in *Computer Networks* (75), pp. 560–577.

Bryman, A., and Bell, E. 2007. *Business Research methods*, Oxford University Press, New York.

Brynjolfsson, E., Hitt, L. 1993. "Is Information Systems Spending Productive? New Evidence and New Results," in *Proceedings of the 14th International Conference on Information Systems*, pp. 47-64.

Cameron, K. S., and Whetten, D. A. 1983. "Some Conclusions about Organizational Effectiveness," in *Organizational Effectiveness: A Comparison of Multiple Models*, K. S. Cameron, and D. A. Whetten (eds.), New York: Academic Press, pp. 261-277.

Chui, M., Manyika, J., Bughin, J., Dobbs, R., Roxburgh, C., Sarrazin, H., Sands, G., and Westergren, M. 2012. "The social economy: Unlocking Value and Productivity Through Social Technologies," *McKinsey Global Institute Report*.

Davenport, T. H., Harris, J., and Cantrell, S. 2004. "Enterprise Systems and On-Going Process Change," *Business Process Management Journal* (10:1), pp. 16-26.

Davis, F. D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly* (13:3), pp. 319–340.

DeLone W. H., and McLean, E. R. 1992. "Information Systems Success: The Quest for the Dependent Variable," *Information Systems Research* (3:1), pp. 60-95.

DeLone, W. H., and McLean, E. R. 2003. "The DeLone and McLean Model of Information Systems Success: A Ten-Year Update," *Journal of Management Information Systems*, (19:4), pp. 9-30.

Desmarais, M. R., Fiset, L. J.-Y., and Talbi, H. 1997. "Cost-Justifying Electronic Performance Support Systems," *Communications of the ACM* (40:7), pp. 39-48.

DiMicco, J.M., Millen, D.R., Geyer, W., Dugan, C., Brownholtz, B., and M. Muller 2008. "Motivations for social networking at work," *Proceedings of the 11th Conference on Computer Supported Cooperative Work*, San Diego.

Du, H. S., and Wagner, C. 2006. "Weblog Success: Exploring the Role of Technology," *International Journal of Human-Computer Studies* (64:9), pp. 789–798.

Dubin, R. 1978. *Theory building*, The Free Press, NewYork.

Duivenvoorde, G. P. J., Briggs, R. O., Kolfschoten, G. L., and de Vreede, G.-J. 2009. "Towards an Instrument to Measure Successfulness of Collaborative Effort from a Participant Perspective," in *Proceedings of the 42nd Hawaii International Conference on System Sciences*.

Easley, R. F., Devaraj, S., and Crant, M. J. 2003. "Relating Collaborative Technology Use to Teamwork Quality and Performance: An Empirical Analysis," *Journal of Management Information Systems* (19:4), pp. 249-270.

Faraj, S., Jarvenpaa, S. L., and Majchrzak, A. 2011. "Knowledge collaboration in online communities, " *Organization Science*, (22:5), pp. 1224–1239.

Fitzgerald, G. 1998. "Evaluating Information Systems Projects: A Multidimensional Approach," *Journal of Information Technology* (13:1), pp. 15-27.

Gable, G. G., Sedera, D., and Chan T. 2003. "Enterprise Systems Success: A Measurement Model," in *Proceedings of the 24th International Conference on Information Systems*, pp. 576-591.

Gable, G. G., Sedera, D., and Chan, T. 2008: "Re-conceptualizing Information System Success: the IS-Impact Measurement Model," *Journal of the Association for Information Systems* (9:7), pp. 1-32.

Goes, P. B. 2014. "Design Science Research in Top Information Systems Journals", *MIS Quarterly* (38:1), pp. iii-viii.

Goldkuhl, G. 2004. "Design Theories in Information Systems - a Need for Multi-Grounding," *Journal of Information Technology Theory and Application* (6:2), pp. 59–72.

Gray, P.H., Parise, S. and Iyer, B. 2011. "Innovation Impacts of Using Social Bookmarking Systems," *MIS Quarterly* (35:3), pp. 629–643.

Gregor, S., and Hevner, A. R. 2013. "Positioning and Presenting Design Science Research for Maximum Impact," *MIS Quarterly* (37:2), pp. 337-355.

Gregor, S., and Jones, D. 2007. "The Anatomy of a Design Theory," *Journal of the Association for Information Systems* (8:5), pp. 312-335.

Gregor, S., Müller, O., and Seidel, S. 2013. "Reflection, Abstraction And Theorizing In Design And Development Research" in *Proceedings of the 21th European Conference on Information Systems*, paper 74.

Grover, V., Jeong, S. R., and Segars, A. H. 1996. "Information Systems Effectiveness: The Construct Space and Patterns of Application," *Information and Management* (31:4), pp. 177-191.

Hevner, A. R., March, S. T., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), pp. 75-105.

Hooker, J. N. 2004. "Is Design Theory Possible?," *Journal of Information Technology Theory and Application (*6:2), pp. 73-83.

Hsu, C., and Lin, J. C. 2008. "Acceptance of Blog Usage. The Roles of Technology Acceptance, Social Influence and Knowledge Sharing Motivation," *Information and Management* (45:1), pp. 65–74.

Ifinedo, P. 2006. "Extending the Gable et al. Enterprise Systems Success Measurement Model: a Preliminary Study," *Journal of Information Technology Management* (XVII:1), pp. 14-33.

Iivari, J. 1983. "Contributions to the Theoretical Foundations of Systemeering Research and the Picoco Model," Oulu, Finland: Institute of Data Processing Science, University of Oulu.

Irani, Z., and Love, P. E. D. 2000. "The Propagation of Technology Management Taxonomies for Evaluating Investments in Information Systems," *Journal of Management Information Systems* (17:3), pp. 161-177.

Janhonen, M., and Johanson, J.-E. 2011. "Role of Knowledge Conversion and Social Networks in Team Performance," *International Journal of Information Management* (31:3), pp. 217-225.

Joosten, D., Basten, D., and Mellis, W. 2011. "Measurement of Information System Project Success in Organizations – What Researchers can Learn from Practice," in *Proceedings of the 19th European Conference on Information Systems*, paper 117.

Kane, G. C., Alavi, M., Labianca, G. J., and Borgatti, S. P. 2014. "What's Different About Social Media Networks? A Framework and Research Agenda," *MIS Quarterly* (38:1), pp. 275–304.

Kane, G. C., Fichman, R. G., Gallaugher, J., and Glaser, J. 2009. "Community relations 2.0," *Harvard Business Review* (87:11), pp. 45–50.

Kettles, D., and David, J. S. 2008. "The Business Value of Social Network Technologies: A Framework for Identifying Opportunities for Business Value and an Emerging Research Program," in *Proceedings Americas Conference on Information Systems,* Toronto, paper 379.

Kiron, D., Palmer, D., Phillips, A., and Berkman, R. 2013. "Social Business: Shifting Out of First Gear," *MIT Sloan Management Review*. Research Report.

Kraut, R., E., and Resnick, P., 2012. *Building Successful Online Communities: Evidence-Based Social Design*, Cambridge, Massachusetts Institute of Technology.

Kristensen, K., and Kijl, B., 2010. "Collaborative Performance: Addressing the ROI of Collaboration," *International Journal of e-Collaboration* (6:1), pp. 53-69.

Kronbichler, S. A., Ostermann, H., Staudinger, R. 2010. "A Comparison of ERP-Success Measurement approaches", *Journal of Information Systems and Technology Management* (7:2), pp. 281-310.

Kurian, D., Galupe R. B., and Diaz, J. 2000. "Taking Stock: Measuring Information Systems Success," *Proceedings of ASAC-IFSAM-2000 Conference*, Montréal, Canada.

Kügler, M., and Smolnik, S. 2013. "Just for the fun of it? Towards a Model for Assessing the Individual Benefits of Employees' Enterprise Social Software Usage," in *Proceedings of the 46th Hawaii International Conference on System Sciences*, pp. 3614-3623.

Kügler, M., Smolnik, S., and Raeth, P. 2013. "Determining the Factors Influencing Enterprise Social Software Usage: Development of a Measurement Instrument for Empirical Assessment," in *Proceedings of the 46th Hawaii International Conference on System Sciences*, pp. 3635-3644.

Li, H., and Gupta, A. 2009. "Understanding the Impact of Instant Messaging (IM) on Subjective Task Complexity and User Satisfaction," in *Proceedings of the Pacific Asia Conference on Information Systems*, paper 10.

Lin, H.-F., and Lee, G.-G. 2006. "Determinants of Success for Online Communities: An Empirical Study," *Behaviour & Information Technology* (25:6), pp. 479-488.

Majchrzak, A., Wagner, C., and Yates, D. 2006. "Corporate Wiki Users – Result of a Survey," in: *WikiSym 2006 - The International Symposium on Wikis*, Odense, Denmark.

Majumdar, A., Krishna, S., and Bjorn, P. 2013. "Managers' Perceptions of Social Software Use in the Workplace: Identifying the Benefits of Social Software and Emerging Patterns of its Use," in *Proceedings of the Nineteenth Americas Conference on Information Systems*, Chicago, Illinois.

March, S. T., and Smith, G. F. 1995. "Design and Natural Science Research on Information Technology," *Decision Support Systems* (15:4), pp. 251- 266.

Markus, M. L., Majchrzak, A. and Gasser, L. 2002. "A Design Theory for Systems That Support Emergent Knolwedge Processes," *MIS Quarterly* (26:3), pp. 179-212.

Markus, M. L., and Tanis, C. 2000. "The Enterprise System Experience – From Adoption to Success," in *Framing the Domains of IT Management: Projecting the Future Through the Past*, Zmud, R.W. (ed.), Cincinnatti: Pinnaflex Educational Resources, Inc. (Chapter 10), pp. 173-207.

Mayring, P. 2000. "Qualitative Content Analysis," *Forum Quality Social Research* (2:1).

McAfee, A. 2009. "Shattering the Myths About Enterprise 2.0," *Harvard Business Review* (78:11), pp. 1-6.

Melville, N., Kraemer, K., and Gurbaxani, V. 2004. "Information Technology and Organizational Performance: An Integrative Model of IT Business Value," *MIS Quarterly* (28:2), pp. 283-322.

Muller, M., J., Freyne, J., Dugan, C., Millen, D. R., and Thom-Santelli, J. 2009. "Return on Contribution (ROC): A metric for Enterprise Social Software," in *Proceedings of the European Conference on Computer Supported Cooperative Work*, Springer, London, pp. 143-150.

Offermann, P., Blom, S., Levina, O., and Bub, U. 2010. "Proposal for Components of Method Design Theories," *Business & Information Systems Engineering* (2:5), pp. 287-297.

Peffers, K., Rothenberger, M., and Kuechler, B. (eds.) 2012. *Design Science Research in Information Systems: Advances in Theory and Practice*, 7th International Conference DESRIST, Las Vegas, NV, USA.

Pirkkalainen, H., Manathunga, A., Pawlowski, J., and Valtonen, J. 2009. "An Evaluation Framework for Social Software Supporting Global Work Processes," *Research Forum to Understand Business in Knowledge Society*.

Raeth, P., Kügler, M., and Smolnik, S. 2012. "The Impact of Organizational Social Web Site Usage on Work Performance: A Multilevel Structural Interaction Perspective," in *Proceedings of the 45th Hawaii International Conference on System Sciences*, pp. 4011 – 4020.

Regmi, K., Naidoo, J., and Pilkington, P. 2010. "Understanding the Processes of Translation and Transliteration in Qualitative research." *International Journal of Qualitative Methods* (9:1), pp. 16-26.

Richter, A., Heidemann, J., Klier, M., and Behrendt, S. 2013. "Success Measurement of Enterprise Social Networks," in *Proceedings 11th International Conference on Wirtschaftsinformatik*, Leipzig, Germany.

Richter, A., & Riemer, K. 2009. "Corporate Social Networking Sites - Modes of Use and Appropriation through Co-Evolution," in *Proceedings of the Australian Conference on Information Systems*, paper 34.

Richter, A., and Riemer, K. 2013a. "Malleable End-user Software," *Business & Information Systems Engineering* (5:3), pp. 195-197.

Richter, A., and Riemer, K. 2013b. "The Contextual Nature of Enterprise Social Networking: A multi case Study Comparison," in *Proceedings of the 21st European Conference on Information Systems*, Utrecht, Netherlands, paper 94.

Rosemann, M., and Vessey, I. 2005. "Linking Theory and Practice: Performing a Reality Check on a Model of IS Success," in *Proceedings of the 13th European Conference on Information Systems (ECIS)*, pp. 854-865.

Rosemann, M., and Wiese, J. 1999. "Measuring the Performance of ERP Software – a Balanced Scorecard Approach," in *Proceedings of the 10th Australasian Conference on Information Systems*, pp. 773-784.

Sabherwal, R., Jeyaraj, A., and Chowa, C. 2006. "Information System Success: Individual and Organizational Determinants," *Management Science* (52:12), pp. 1849-1864.

Schryen, G. 2010. "Preserving Knowledge on IS Business Value. What Literature Reviews Have Done," *Business & Information Systems Engineering* (2:4), pp. 225-237.

Schubert, P., and Williams, S. P. 2011. "A Framework for Identifying and Understanding Enterprise Systems Benefits," in *Business Process Management Journal* (17:5), pp. 808-828.

Schubert, P., and Williams, S. P. 2013a. "Management der Nutzenrealisierung aus Informationstechnologie," in *Wirtschaftsinformatik Proceedings*, paper 38.

Schubert, P., and Williams, S., P. 2013b. "The Concept of Social Business: Oxymoron or Sign of a Changing Work Culture?," in *Proceedings of the 26th Bled Conference*, 2013.

Schultze, U., and Avital, M. 2011. "Designing Interviews to Generate Rich Data for Information Systems Research," *Information and Organization* (21:1), pp. 1-16.

Seddon, P.B., Calvert, C., and Yang, S. 2010. "A Multi-project Model of Key Factors Affecting Organizational Benefits from Enterprise Systems," *MIS Quarterly* (34:2), pp. 305-328.

Seddon P. B., Staples S., Patnayakuni R., Bowtell M. 1999. "Dimensions of Information Systems Success," *Communications of the Association for Information Systems* (2:20).

Shang, S., and Seddon, P. B. 2002. "A Comprehensive Framework for Assessing and Managing the Benefits of Enterprise Systems: The Business Manager's Perspective," *Information Systems Journal* (12:4), pp. 271-299.

Simon, H. 1996. *The Sciences of the Artificial, 3rd edition*, Cambridge, MA: MIT Press.

Smits, M., and Mogos, S. 2013. "The Impact of Social Media on Business Performance," in *Proceedings of the 21st European Conference on Information Systems*. Paper 713.

Soh, C., and Markus, M. L. 1995. "How IT Creates Business Value: A Process Theory Synthesis", in *Proceedings of the 16th International Conference on Information Systems,* paper 4.

Spagnoletti, P., and Tarantino, L. 2013. "User Centered Systems Design: The Bridging Role of Justificatory Knowledge," in *Designing Organizational Systems*, Baskerville, R. et al. (eds.) Heidelberg: Springer-Verlag, pp. 105-121.

Spencer, L., Ritchie, J., Lewis, J., and Dillon, J. 2003. "Quality in Qualitative Evaluation: A Framework for Assessing Research Evidence," National Centre for Social Research, Government Chief Social Researcher's Office, London.

Stefanou, C. J. 2001. "A Framework for the Ex-ante Evaluation of ERP Software," in *European Journal of Information Systems* (10:7), pp. 204–215.

Steinhüser, M., Richter, A., and Smolnik, S. 2015. "How to bridge the boundary? Determinants of inter-organizational social software usage," in *Electronic Markets - The International Journal on Networked Business*.

Steinhüser, M., Smolnik, S., Hoppe, U. 2011. "Towards a Measurement Model of Corporate Social Software Success – Evidences from an Exploratory Multiple Case Study," in *Proceedings of the Hawaii International Conference on System Sciences*, Kauai: USA.

Stocker, A., Richter, A., Hoefler, P. & Tochtermann, K. 2012. "Exploring Appropriation of Enterprise Wikis: A Multiple-Case Study," in *Computer Supported Cooperative Work: The Journal of Collaborative Computing* (21:2), pp 317-356.

Stockmann, R. 2011. *A Practitioner Handbook On Evaluation*, Cheltenham, UK and Northampton, USA: Edward Elgar Publishing, pp. 13-29.

Tallon, P. P., Kraemer, K. L., and Gurbaxani, V. 2000. "Executives' Perceptions of the Business Value of Information Technology: a Process-oriented Approach," in *Journal of Management Information Systems* (16:4), pp. 145–173.

Tarafdar, M., Tu, Q., and Ragu-Nathan, T., S. 2010. "Impact of Technostress on End-User Satisfaction and Performance," *Journal of Management Information Systems (*27: 3), pp. 303–334.

Trimi, S., and Galanxhi-Janaqi, H. 2008. "Organisation and Employee Congruence: A Framework for Assessing the Success of Organisational Blogs," *International Journal of Information Technology & Management* (7:2), pp. 120–133.

Urbach, N., Smolnik, S., and Riempp, G. 2009a. "The State of Research on Information Systems Success - a Review of Existing Multidimensional Approaches," *Business & Information Systems Engineering* (1:4), pp. 315-325.

Urbach, N.; Smolnik, S.; Riempp, G. 2009b. „Development and Validation of a Model for Assessing the Success of Employee Portals," *Proceedings of the 17th European Conference on Information Systems*, paper 21.

Venable, J. 2006. "The Role of Theory and Theorising in Design Science Research," in *Proceedings of the first International Conference on Design Science Research in Information Systems and Technology*, pp. 1-18.

Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. 2003. "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly* (27:3), pp. 425–478.

Walls, J. G., Widemeyer, G. R., and El Sawy, O. A. 1992. "Building an Information System Design theory for Vigilant EIS," *Information Systems Research*, (3:1), pp. 36-59.

Walsham, G. (2006). "Doing interpretive research," *European Journal of Information Systems*, (15:3), pp. 320–330.

Webster, J., and Watson, R. T. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review," *MIS Quarterly* (26:2), pp. xiii-xxiii.

Weill, P. 1992. "The Relationship Between Investment in Information Technology and Firm Performance: a Study of the Valve Manufacturing Sector," *Information Systems Research* (3:4), pp. 307-333.