Open

Research Article

Process is king: Evaluating the performance of technology-mediated learning in vocational software training

Matthias Söllner[1,2], Philipp Bitzer[2], Andreas Janson[2],
Jan Marco Leimeister[1,2]

[1]Institute of Information Management, University of St.Gallen, Müller-Friedberg-Strasse 8, 9000 St.Gallen, Switzerland;
[2]Information Systems, Research Center for IS Design (ITeG), Information Systems, University of Kassel, Pfannkuchstr. 1, Kassel 34121, Germany

Correspondence:
M Söllner, Institute of Information Management, University of St.Gallen, Müller-Friedberg-Strasse 8, 9000 St.Gallen, Switzerland.
Tel: +41 (0)71 224 3838;
Fax: +41 (0)71 224 3296
E-mail: matthias.soellner@unisg.ch

Abstract

Technology-mediated learning (TML) is a major trend in education, since it allows to integrate the strengths of traditional- and IT-based learning activities. However, TML providers still struggle in identifying areas for improvement in their TML offerings. One reason for their struggles is inconsistencies in the literature regarding drivers of TML performance. Prior research suggests that these inconsistencies in TML literature might stem from neglecting the importance of considering the process perspective in addition to the input and outcome perspectives. This gap needs to be addressed to better understand the different drivers of the performance of TML scenarios. Filling this gap would further support TML providers with more precise guidance on how to (re-)design their offerings toward their customers' needs. To achieve our goal, we combine qualitative and quantitative methods to develop and evaluate a holistic model for assessing TML performance. In particular, we consolidate the body of literature, followed by a focus group workshop and a Q-sorting exercise with TML practitioners, and an empirical pre-study to develop and initially test our research model. Afterward, we collect data from 161 participants of TML vocational software trainings and evaluate our holistic model for assessing TML performance. The results provide empirical evidence for the importance of the TML process quality dimension as suggested in prior literature and highlighted by our TML practitioners. Our main theoretical as well as practical contribution is a holistic model that provides comprehensive insights into which constructs and facets shape the performance of TML in vocational software trainings.
Journal of Information Technology (2018) **33**, 233–253. https://doi.org/10.1057/s41265-017-0046-6; published online 26 June 2017

Keywords: technology-mediated learning; holistic evaluation of performance; perceived TML process quality; vocational software trainings

The online version of this article is available Open Access

Introduction

In 2010, more than 70% of all German companies invested in vocational training and 94% of all major companies (more than 1000 employees) invested in according trainings (Vollmar, 2013). This accounts for a market volume of more than 28bn Euros (Seyda and Werner, 2012). These numbers reflect not only the major economic importance of such

learning services, but also the need to constantly train employees in order to remain competitive and avoid the loss of knowledge (Vollmar, 2013). In this context, technology has a major influence on all learning scenarios (Gupta and Bostrom, 2009), therefore indicating that technology-mediated learning (TML) is a major trend in education. TML offers the opportunity to integrate the strengths of synchronous (face-to-face) and asynchronous (IT-based) learning activities (Garrison and Kanuka, 2004). Thus, the importance of TML is about to further increase, since it empowers the design of innovative, more individual, and resource-preserving ways of learning, for example, micro-learning at the workplace or location-independent, cloud-based learning.

Despite their many advantages, such as improved economic feasibility (Wegener *et al.*, 2012) or improved student achievements (Alonso *et al.*, 2011; López-Pérez *et al.*, 2011), TML poses several fundamental challenges related to their variability, i.e., research struggles to fully understand the effects of synchronous and asynchronous learning elements on TML outcomes due to the complexity of TML. Among others, Bitzer and Janson (2014) identified a broad range of indicators with inconsistent effects in their extensive literature review, in addition to similar results in prior literature reviews on TML (Gupta and Bostrom, 2009; Gupta *et al.*, 2010). This is especially challenging for providers of respective vocational trainings using TML for their offerings, since they lack reliable approaches to evaluate the performance of their offerings and to derive respective improvements. Gupta and Bostrom (2009) highlight that one probable reason for these inconsistencies is the fact that many studies use input–outcome research designs that ignore critical aspects of the learning process. This criticism especially refers to research designs that merely consider the impact of learning methods on learning outcomes without considering psychological processes and factors of the learning process dimension (Alavi and Leidner, 2001a). Therefore, a holistic evaluation of TML performance is essential, taking into account not only selected elements and their effects on the TML outcomes, but also the learning process in TML, that is, the learning process of a learner that participates in TML scenarios (Hattie and Yates, 2014). Consequently, there is a research gap in terms of a comprehensive explanation of causal relationships within TML scenarios to ensure a correct performance evaluation, allowing to derive general, transferable advice for the design of TML scenarios (Alavi and Leidner, 2001b).

Hence, the aim of our study is to develop a comprehensive approach for investigating the performance of TML, especially considering the effects of input-, process-, and outcome-related constructs. In particular, we aim to answer the following research questions:

1. Which constructs and facets should be included in a comprehensive model for evaluating the performance of TML?
2. What impacts do constructs related to TML inputs and the TML process have on each other and on constructs resembling TML outcomes?

Our synthesis of the existing literature combined with the expertise of 12 experts from providers of vocational software trainings resulted in two input-, one process-, and two outcome-related constructs that should be included in a holistic evaluation approach concerning TML performance. The main contribution of our study is the development and evaluation of a holistic model for assessing the performance of TML that accounts for the fact that input-, process-, and outcome-related constructs should be considered. In addition, we take a deeper look into the most important facets shaping the input- and process-related constructs by conducting a rigorous scale development process with experts working for vocational training providers. Furthermore, we examine the interplay of the different constructs of the model, allowing deeper insight into the structural relationships relevant in the context of evaluating the performance of TML. Following the terminology of Gregor (2006), we provide a theory of explanation and prediction for TML performance that explicitly considers input-, process-, and outcome-related constructs, as well as the facets shaping them.

To achieve our desired research objective, the remainder of this paper is structured as follows: First, we present related work regarding research on TML. Based on these foundations, we develop our research model and derive our hypotheses. After explaining our research method, which encompasses a literature review, a focus group workshop, a Q-sort application, and the partial least squares (PLS) approach applied in our pre- and main study, we present our results and discuss their implications for theory and practice. The paper closes with the limitations of our study and areas for future research.

Theoretical background

Technology-mediated learning

As already mentioned in the introduction, technology-mediated learning (TML) offers variable learning scenarios for training purposes. In a comprehensive sense, TML describes "environments in which the learner's interactions with learning materials (readings, assignments, exercises, etc.), peers, and/or instructors are mediated through advanced information technologies" (Alavi and Leidner, 2001a: 2). In consequence, research often uses the term e-learning as a synonym (Gupta and Bostrom, 2013). However, it should be noted that TML has many variations in practice and often constitutes a combination of different learning modes and methods and can therefore be considered a blended learning service. Such a blended approach could be designed with the following elements according to Gupta and Bostrom (2009):

- Web- or computer-based approaches,
- Asynchronous or synchronous,
- Led by an instructor or self-paced,
- Individual- or team-based learning modes.

This variety of possible combinations poses many challenges to TML research. In consequence, empirical TML research has found mixed results concerning the impact of TML that are related to the individual as well as the team level (Gupta and Bostrom, 2009). One possible explanation is the focus of TML studies on input–outcome research designs that consider the above-listed elements of TML but neglect, among other things, the learning process (Alavi and Leidner, 2001a; Hannafin *et al.*, 2004). This is highlighted by our review of TML research regarding insights into the specific TML dimensions and their according findings (see Table 1, based on the findings of Bitzer and Janson, 2014).

Considering previous studies, seminal research regarding TML has focused on the effects of the structural potentials of IT-supported collaborative learning, especially with the

Table 1 Research directions for TML

Topic	Key Findings and Exemplary References	Research Directions
Predisposition Quality	General attitudes toward TML (Arbaugh, 2001), computer experience (Arbaugh, 2013), self-efficacy (Lim *et al.*, 2007; Eom, 2015; Eom *et al.*, 2006), learning styles (Ozkan and Koseler, 2009), and learning motivation (Klein *et al.*, 2006) influence learning outcomes positively	How do individual differences that determine predisposition quality influence the relationship between provided structural potentials, the learning process, and learning outcomes?
Structural Quality	System (Alavi *et al.*, 2002; Lin, 2007) and information quality (Eom, 2011) as well as trainer characteristics (Sun *et al.*, 2008) influence learning outcomes positively	How does structural quality influence the learning process and how does this relate to learning outcomes?
TML Process Quality	Process determinants, interaction (Siau *et al.*, 2006; Arbaugh, 2000), and support (Ozkan and Koseler, 2009) influence TML outcomes positively	What are determinants of the TML process and how does TML process quality relate to TML outcomes and its structural antecedents?
Outcome Quality	TML outcomes are significantly influenced by structural and process quality (Gupta and Bostrom, 2013)	How do structural potentials and TML process quality influence the quality of TML outcomes?

deployment of group support systems for the purpose of learning (Alavi, 1994; Alavi *et al.*, 1995, 1997). In the early 2000s, research shifted to a more learning process-centered view (Alavi *et al.*, 2002; Alavi and Leidner, 2001a) that is supported by recent empirical research regarding the influence of the learning process on learning outcomes (e.g., Gupta and Bostrom, 2013). However, until today, research has not considered an integrative TML assessment to evaluate TML quality from a holistic perspective, including input-, process-, and outcome-related constructs. Such an approach is necessary for numerous reasons. First, the learning process is the core construct that acts as a mediator of TML input dimensions and therefore influences the outcomes of TML. Second, without such a comprehensive assessment of the TML performance, neglecting, for instance, the learning process quality might lead to inconclusive results (Gupta and Bostrom, 2009). Third, previous studies are limited to a certain technology or learning method and their effects on training outcomes, which prevents to acquire insights transferable to other TML scenarios. In sum, a comprehensive TML assessment has to be applicable for a wide range of blended learning services to achieve comparable TML performance assessments that are not limited to a particular context. To achieve such a comprehensive TML assessment, it is necessary to be aware of the different facets of the different constructs, which we will identify in the following section.

Identification of facets of core constructs in the context of TML[1]
In addition to the theoretical background of TML, a deeper view into the facets of the specific TML core constructs is needed to understand which facets form constructs and to ensure a proper measurement of the constructs essential for our study. In the field of TML, a varying amount of research has been conducted concerning the various dimensions. Predisposition quality has been intensively examined, and learner characteristics such as (meta-) cognition (Pintrich and De Groot, 1990), motivation (Cole *et al.*, 2004; Colquitt *et al.*, 2000; Pintrich and De Groot, 1990), self-efficacy/ learning management (Colquitt *et al.*, 2000; Tannenbaum *et al.*, 1991), and technology readiness (van der Rhee *et al.*, 2007) were shown to play an important role for the TML process and outcome (Gupta *et al.*, 2010).

Regarding the structural quality of TML, various aspects have been used to describe the structural potential for the provision of TML. The IT system quality of the training software and the applied e-learning tools determine the perceived process quality and the corresponding outcome quality (DeLone and McLean, 2003; Lin, 2007; Petter *et al.*, 2008). Further, the information quality and the quality of the learning materials, respectively, determine the outcome quality (Petter *et al.*, 2008; Ozkan and Koseler, 2009; DeLone and McLean, 1992; Rasch and Schnotz, 2009). The trainer characteristics can be divided into the following aspects. First, the didactical competence of the trainer can be considered an important determinant for TML success (Arbaugh, 2001; Armstrong *et al.*, 2004; Kim *et al.*, 2011). The professional competence also influences TML (Jacobs *et al.*, 2011; Ozkan and Koseler, 2009; Schank, 2005). Further, social skills such as attitude toward the students (Choi *et al.*, 2007) play a decisive role in TML scenarios. Finally, the learning environment, that is, the classroom or the virtual learning environment, should be considered for a comprehensive TML evaluation.

Also regarding TML outcomes, a plethora of different research insights can be found. Kirkpatrick and Kirkpatrick (2005) suggest the use of four learning dimensions when considering the systematic evaluation of vocational training and education: Reaction describes the emotional reaction to the course, learning refers to the learning success, application of knowledge relates to the actual usage of acquired knowledge in the real world, and company success describes effects within the company caused by the knowledge of the course participant. The most common measure for reaction is satisfaction, one of the frequent measures used in the evaluation of TML, also considered as an affective learning outcome measure (Alonso *et al.*, 2011; Arbaugh, 2001; Johnson *et al.*, 2009; Gupta *et al.*, 2010). Another frequent measure that is examined is learning success, often referred to as course performance (Arbaugh, 2001; Benbunan-Fich and Arbaugh, 2006; Hiltz *et al.*, 2000; Santhanam *et al.*, 2008). This measure can, on the one hand, be considered as a subjective measure of perceived learning success in accordance with Alavi (1994), which is a meta-cognitive measure of learning outcomes. On the other hand, when considering objective measures, learning success

can be considered as cognitive knowledge acquisition related, for instance, to declarative or procedural knowledge. Several authors include the application of knowledge (Hansen, 2008; Sousa *et al.*, 2010) as a skill-based outcome referring to a behavioral change, for example, if an individual exhibits a better (delayed) task performance (Yi and Davis, 2003). However, only a few consider business effects of training (Aragon-Sanchez *et al.*, 2003; Reber and Wallin, 1984).

However, regarding TML process quality, only few insights can be found in prior research. While we could draw on existing and established measures in educational literature, for example related to approaches how someone learns (see, e.g., Biggs, 1993), such measures are with limited relevance for our facet identification, since they are not giving systematic advice how to actually improve the perceived quality of the TML process. Rather, measurement approaches from educational literature are focusing on the individual and their cognitive and motivational processes. This is indeed very important to describe how individuals differ in learning, how tasks are handled and even how teaching contexts differ (Biggs *et al.*, 2001), but neglects how teaching contexts, for example, in TML should be designed to deliver a high quality. Concerning TML research in specific, interactivity is the most profoundly studied process-specific component that has been the focus of research (Bitzer and Janson, 2014). Interaction and interactivity, including interaction among learners, learner–lecturer interaction, and learner–IT interaction (Brower, 2003; Cole *et al.*, 2004; Evans and Gibbons, 2007; Lehmann *et al.*, 2016; Sims, 2003; Smith and Woody, 2000; Thurmond and Wambach, 2004; Moore, 1989), have been widely examined and can be considered an important facet of TML success. In addition, just recently, the appropriation of training methods has been examined, providing insights into the facets of this construct, such as technology mediation or collaboration (Gupta and Bostrom, 2013).

To summarize, whereas many insights into the facets of four of our five constructs can be found in the literature, perceived TML process quality has received only little attention until now. Transfer efficiency, namely a faster and efficient way of working, and effects on retention seem to be especially highly relevant for measuring the productivity of TML. User satisfaction, even with the use of IT systems and system usability, is a pivotal factor for the success of IS (Wang *et al.*, 2007), but is rare in recent, relevant literature and not substantial enough in the context of perceived TML process quality. As a result, before being able to reliably and validly evaluate our hypotheses, which are derived in the following section, we decided to conduct a scale development process, as described below, to (a) consolidate the insights found in literature, and (b) create further insights, particularly into the facets of perceived TML process quality, using TML experts. This procedure helps us to ensure the proper measurement of our constructs and thus increases the reliability and validity of our results. Further details can be found in the research method section after the development of our hypotheses.

Hypotheses development

For our overall research model, we will derive according to the hypotheses in this section. The hypotheses of the model are on a construct level, concentrating on the identified constructs and their interplay in a holistic TML evaluation model.

Following the foundations laid in the previous section, the individual differences of learners influence the TML process through complex interaction of the learner with the learning methods and structures (Gupta *et al.*, 2010). For example, attitudes toward technology (van der Rhee *et al.*, 2007) or their ability to organize their learning activities (Colquitt *et al.*, 2000) are supposed to have an impact on the TML process, since a higher predisposition quality ensures that learners are more content in their TML processes. Furthermore, predisposition quality affects how learners in TML pay attention to learning methods and structures, and how they process these in terms of effectiveness (Pintrich and De Groot, 1990; Deci and Ryan, 2000). For example, a high predisposition quality enables the participants to persist at a difficult task, to fade out distractors, and thus likely has a positive impact on the perceived quality of the TML process, since the participants can focus on and follow the intended TML process more effectively. Therefore, our first hypothesis is:

> **H1** Predisposition quality has a positive impact on the perceived quality of the TML process.

Furthermore, the perceived structural quality provided by the TML provider determines the perceived quality of the TML process. We refer to the effect of the quality of provided structures when designed in accordance with the learning process and induced by the overall epistemological perspective of TML (Gupta and Bostrom, 2009). An example is the case of TML that is designed in accordance with a cooperative learning model (Leidner and Jarvenpaa, 1995) that should therefore engage collaborative learning processes among learners. Consequently, high-quality collaborative learning structures for TML ensure successful collaborative learning processes. As a result, structural quality is supposed to positively influence the TML process. Furthermore, insights from cognitive load theory suggest that people can only process a certain amount of information at a time (Sweller and Chandler, 1991). If a TML provider is able to provide structures that are of high quality in terms of inducing lower cognitive load, the participants will be less distracted and can thus focus on the content and the learning process. At the same time, if the structures provided are of poor quality, participants might, for example, wonder what they are currently about to do or which learning goal they are supposed to reach. Thus, they are distracted from acquiring knowledge, since the limited cognitive resources are used by individual learners to comprehend the design of the learning material and not the learning content itself. This distraction will then likely to result in a lower TML process quality. This leads to our second hypothesis:

> **H2** Perceived structural quality has a positive impact on the perceived quality of the TML process.

Prior literature has shown that the cognitive aspects of learners' activities influence TML outcomes directly through the formation of mental models (Gupta *et al.*, 2010). When considering predisposition factors such as attitudes toward IT, learners with a more positive attitude might not be afraid of using IT in TML, which will result in higher levels of satisfaction and perceived learning success (Piccoli *et al.*, 2001; Sun *et al.*, 2008). Moreover, factors such as cognition and motivation are known for their positive influence on

learning success (Pintrich and De Groot, 1990). In the same vein as hypothesized in H1, learners with, for example, high motivational levels (Linnenbrink and Pintrich, 2002; Deci and Ryan, 2000) and therefore a high predisposition quality are likely to exhibit a more goal-directed behavior as well as a more effective cognitive processing when participating in TML, thus resulting in higher satisfaction and perceived learning success. Therefore, our next two hypotheses constitute:

H3 Predisposition quality has a positive impact on the perceived learning success of the TML participants.

H4 Predisposition quality has a positive impact on the satisfaction of the TML participants.

Besides its already hypothesized positive impact on the TML process, perceived structural quality is also supposed to directly influence the TML outcomes, that is, learning success and satisfaction (Kirkpatrick and Kirkpatrick, 2005). When the provided structural components support the underlying epistemological perspective and are of high quality, learning outcomes are influenced positively (Gupta and Bostrom, 2009). This is, for example, highlighted by the study by Gupta and Bostrom (2013), who investigated how higher levels of enactive and vicarious learning relate to the satisfaction and success of learning. Related to our context, high perceived structural quality provided by the TML provider should have a positive impact on the TML outcomes. For example, well-structured learning materials will support the participants in focusing on understanding the relevant content of the training, thus fostering the participants perceived learning success and satisfaction with the training. Low perceived structural quality, e.g., a demotivated trainer or unstructured learning materials, will hinder the participants in focusing their attention on the content of the training, and thus limiting their perceived learning success as well as their satisfaction with the training. As a result, we can derive two further hypotheses.

H5 Perceived structural quality has a positive impact on the perceived learning success of the TML participants.

H6 Perceived structural quality has a positive impact on the satisfaction of the TML participants.

Regarding the relationship between perceived TML process quality and TML outcomes, the adaptation of the prepared structural elements, for example, learning materials and the training concept, by the learner during the learning process needs to be considered (Gupta and Bostrom, 2009). Therefore, recent research has started to consider the learning process by means of analyzing the procedural factors of TML, focusing on the interaction between learners and the structural potential of IT-mediated learning (Gupta and Bostrom, 2013; Bitzer *et al.*, 2013; Bitzer and Janson, 2014). In this context, the learning process is a complex phenomenon including cognitive processes and interactions based on the aforementioned learning methods, individual differences between the learners, and other elements of the teaching/learning scenarios that influence TML outcomes (Gupta *et al.*, 2010; Gupta and Bostrom, 2013). Based on prior research highlighting the importance of the learning process, we argue that a high-quality TML process is likely to have several positive effects for TML participants.

Consider, for example, the typical case of a software training course that is supplemented by means of a learning management system to structure the TML process. If the quality of interaction and IT support with this learning management system does not fit the expectations of the participants, for example when IT support in the TML process is very low and the participants are overwhelmed by lots of additional learning material provided in the learning management system which they perceived to only add little value, TML results may be endangered. As a result, the participants might perceive TML as unsatisfactory, since they cannot acknowledge how the technology adds value to the learning process. Also, the participants might even ignore relevant learning materials provided in a low-quality TML process, ultimately resulting in lower perceived learning success. However, if we assume that the TML provider ensures a high-quality process guiding the participants through the context, and helping them, for example, to also apply and deepen their newly acquired knowledge, the participants are likely to have a higher perceived learning success, and be more satisfied with the training, e.g., since they recognize that the TML provider has put a lot of effort in carefully designing the TML process to ensure a high-quality experience of the participants. Thus, we derive the following two hypotheses:

H7 Perceived TML process quality has a positive impact on the perceived learning success of the TML participants.

H8 Perceived TML process quality has a positive impact on the satisfaction of the TML participants.

Furthermore, we examine the relationship between perceived learning success and satisfaction. In the literature, learning satisfaction is often used as a proxy for learning success (Sun *et al.*, 2008; Wang, 2003), implying that learning success results in learner satisfaction. In customer satisfaction theory, a comparable observation was made. Here, the perceived performance of a product influences how satisfied the customer is with the product he or she bought (e.g., Tse and Wilton, 1988). When keeping in mind that the learners participate in TML to learn, for example, basics of enterprise resource planning (ERP) and how to operate an ERP software, then their perception of how well they achieved this goal will influence how satisfied they are with the training. If they believe that they learned a lot, they will be satisfied with the TML, and at the same time, they will not be satisfied if they have the feeling that they did not learn anything new or useful. Thus, our last hypothesis is:

H9 Perceived learning success has a positive impact on the satisfaction of the TML participants.

In sum, we derived nine hypotheses (see Figure 1 for a graphical illustration of our research model) that aim to identify causal relationships between the five constructs necessary for a holistic evaluation of TML. To achieve this holistic TML performance evaluation, we now identify the key facets of the previously introduced TML dimensions with a rigorous scale development process in the next section.

Research method

Analogous to the established scale development and validity guidelines (Churchill Jr, 1979; Straub, 1989), we conducted a three-step process to ensure the quality of our scales before
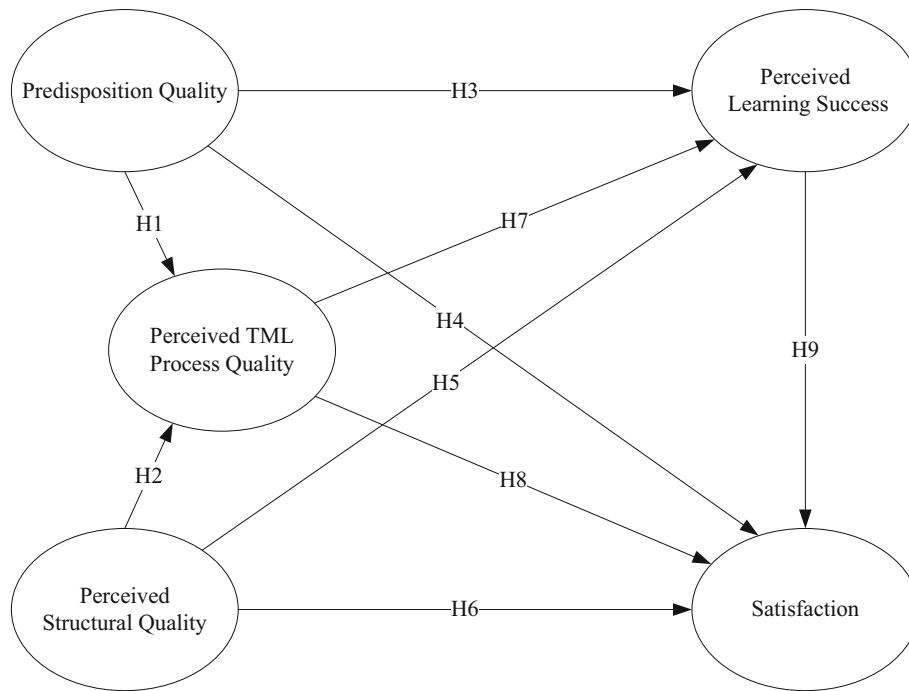
Figure 1 Research model.

finally conducting our main study. We employed the plurality of methods, such as a literature review (1) and a focus group workshop (1b), for the conceptual development of our scales. Next, we completed a conceptual refinement, applying the Q-sort method (2). We then carried out an empirical pre-study with 163 students (3). After the scale development procedures, we conducted our empirical main study with 161 participants of vocational software trainings (4) to evaluate our hypotheses. The research process is illustrated in Figure 2.

Step 1: conceptual development
In a first step, we identified articles and corresponding results regarding factors that influence non-IT-related and IT-related learning scenarios in terms of the four dimensions: predisposition quality, structural quality, TML process quality, and outcome quality. In preparation for the latter quantitative study, we started with an extensive literature review (please see Bitzer and Janson, 2014 for further details).

Based on the results of the literature review (the results are presented in the subsection *Identification of Facets of Core Constructs in the Context of TML* of the theoretical background of this paper), an expert focus group workshop was conducted in an eight-hour setting with twelve experts of the subject matter on hand. The workshop participants were lecturers from an educational background with a minimum of eight years of professional experience. Following the focus group design approach by Frey and Fontana (1991), we designed the focus group taking into account data-related design requirements, interviewers, and group characteristics. Returning to the results of the literature review presented previously, a conceptual model for TML evaluation was derived, including a set of possible categories and corresponding items. The focus group findings were the

exploratory groundwork to add to the existing findings in the literature for the quantitative evaluation and improvement of the existing findings (Miles and Huberman, 1994).

In accordance with Kolfschoten *et al.* (2006), we invited experts to brainstorm about drivers and outcome factors in an initial brainstorming session. This session was followed by an organizational activity based on the expert and literature results that aimed to clarify the existing influencing factors, before the results were finally organized in corresponding dimensions. We thereby developed an initial conceptual model including a total of 24 constructs and 146 corresponding items.

Step 2: conceptual refinement
Next, we applied the Q-sort method to ensure reliability and construct validity of the questionnaire items (Nahm *et al.*, 2002). We asked four experts with more than four years of experience in TML to sort every item according to the identified components in order to improve the comprehensibility and clarity of the items and components. Each expert was presented with the components, corresponding definitions, and a bucket of randomly sorted items printed on small notes. The experts had to assign each note individually to one of the components; alternatively, in case of doubt, notes were sorted into an "unclear" bucket. We interviewed the experts as to why they were unsure about certain items and collected the according feedback. First, we asked two experts to sort the items, after which their feedback was collected and the questionnaire improved. Next, we asked another two experts to conduct the Q-sort. After this procedure was completed, Cohen's kappa, a measure of agreement, exceeded 0.76, representing an excellent degree of agreement beyond chance (Landis and Koch, 1977). Moreover, we used the total hit ratio to identify potential problem
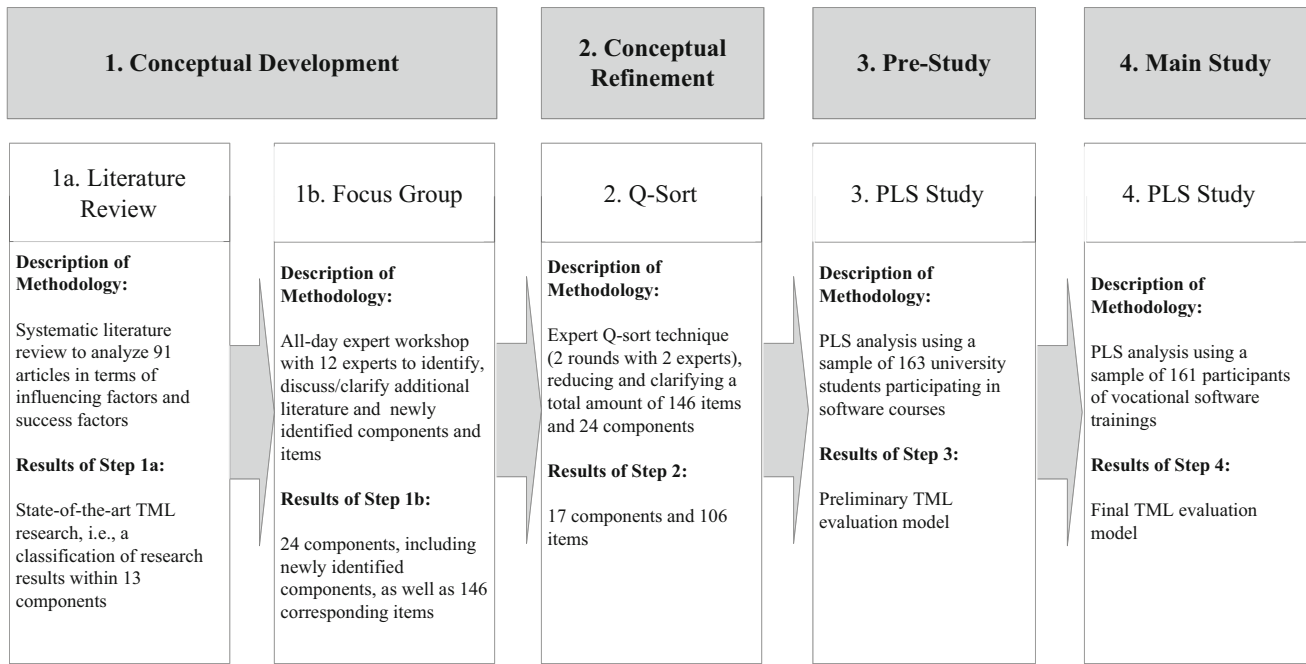
Figure 2 Overview of the Research Process.

component areas. Finally, we were able to eliminate a total of 28 items that at least three of the experts claimed to be irrelevant or unclear. Moreover, we clarified another nine items in terms of wording precision. This step resulted in 17 components and a total number of 106 items.

Step 3: pre-study
The results from steps 1 and 2 served as a foundation for the pre-study (see Figure 2) that was carried out in accordance with the PLS approach (Wold, 1982). To operationalize the results of the first two steps in terms of empirical measurement, we decided to rely on reflective first-order, formative second-order measurement models for predisposition quality, structural quality, and TML process quality, and on reflective measurement models for the two TML outcomes, satisfaction and learning success. In order to evaluate our research model (see Figure 1), we conducted a web-based questionnaire among students who had recently participated in at least one software training. By means of a web-based questionnaire, we accounted for the fact that participants of software training usually face several questionnaires, for example, regarding their satisfaction with the trainer, the course, and so on. Consequently, the willingness to complete another questionnaire during the course could be comparably low, and an ex post assessment using a web-based approach seemed to be a better approach in terms of response rate and data quality. In total, we gathered 163 complete data sets that could be used for our evaluation. The participants were on average 24.67 years of age; 52 of them were female and more than 100 of them were business students. For the data analysis, we relied on the SmartPLS 2.0 M3 software.

Based on the results of our pre-study (for further details, please see Bitzer *et al.*, 2013), the feedback of reviewers and attendees of the International Conference on Information Systems 2013, several dimensions of the predisposition

quality and TML process quality constructs were altered. Regarding predisposition quality, we included the dimension *self-efficacy* to replace *(previous) knowledge*, since this construct better fits our intention to capture the capability of a recipient and not, for example, knowledge about the content of the course. Furthermore, we included the dimension *intrinsic value* as a combination of motivation and expectations. Since the participants of our main study were not students but participants of vocational software trainings, we further included the dimension *perceived importance*, resembling the importance of the training for their job. Furthermore, for the same reason, we added *company support* as a dimension of TML process quality. Based on the results of our pre-study, the constructs and subdimensions presented in Table 2 are used in our main study.

Step 4: main study
Based on the results of the pre-study, we conducted our main study with participants of vocational software trainings. In collaboration with one of our partners, we could access the participants of a series of 10-day vocational software trainings focusing on established ERP software, which were provided by one of the major providers of such trainings in Germany. The participants completed two questionnaires including the items related to our five constructs. The first questionnaire captured predisposition quality, whereas we captured the other four constructs in the second questionnaire. We asked the trainers to hand out the first questionnaire on the first or second day of each course. The second questionnaire was completed on one of the last two days of each course (the complete set of all items used in the main study is found in "Appendix 1"). In total, 161 participants completed both questionnaires, which could be used for our evaluation. The participants were on average 25.57 years of age and 54 were female. For the data analysis, we used the software packages IBM SPSS 21 and SmartPLS 3.0 v3.2.3

Table 2 Constructs and subdimensions used in the main study

Subdimension	Description	Source
Perceived Structural Quality		
Trainer Quality	Trainer characteristics such as professional, didactical, and social competence	Arbaugh (2001) and Kim *et al.* (2011)
Learning Environment	Offline and online environment of the TML	Ladhari (2009)
IT Systems Quality	Quality of the provided IT systems	DeLone and McLean (2003)
Learning Materials	Information quality of the learning materials	DeLone and McLean (2003)
Predisposition Quality		
Self-Efficacy	Recipients' knowledge regarding the course content	Pintrich and De Groot (1990)
Intrinsic Value	Recipients' motivation and expectations	Pintrich and De Groot (1990) and Colquitt *et al.* (2000)
Technology Readiness	Recipients' attitude toward technology	van der Rhee *et al.* (2007)
Self-Regulated Learning	Recipients' capability to organize their own learning activities	Pintrich and De Groot (1990)
Perceived Importance	Recipients' perceived importance of the training	Focus Group
Perceived TML Process Quality		
Interactivity	Interactivity during the course, online and offline	Siau *et al.* (2006)
IT Process Support	Usefulness of the applied IT tools in terms of communication and learning support during the process	Focus Group
Learning Group	Characteristics of the learning group such as homogeneity of knowledge, expectations, or mutual support	Focus Group
Quality of Exercises	Quality of the exercises in terms of usefulness, didactical appropriateness, and understandability	Focus Group
Transparency of the Training Process	Traceability of the course procedure, upcoming process activities, and corresponding learning goals	Focus Group
Company Support	Support through the participants' companies, for example, by releasing them from any further duties during the training	Focus Group
Fit	Overall fit of the course design for the recipient characteristics and expectations	Focus Group
Satisfaction		
No Subdimension	Service recipient satisfaction	Arbaugh (2001)
Perceived Learning Success		
No Subdimension	Service recipient knowledge gain	Arbaugh (2001)

(Ringle *et al.*, 2015). To ensure the correctness of our factor weights, loadings, as well as path coefficients, we applied the PLSc algorithm (Dijkstra and Henseler, 2015). To overcome issues related to missing values in the sample, such as an underestimation of the actual variance or systematical reduction in variable correlations, we applied a multiple imputation approach in order to replace missing values in the data set (Rubin, 1987). To evaluate our reflective first-order, formative second-order measurement models, we used a two-step approach and first computed the factor scores for the reflective first-order factors that were then used as indicators in the second step (Chwelos *et al.*, 2001; Becker *et al.*, 2012; Söllner et al., 2012, 2016; Wang and Benbasat, 2005).

Recently, a number of researchers have brought up the problem of common method variance in behavioral research (Podsakoff *et al.*, 2003; Sharma *et al.*, 2009). These publications point out that a significant amount of variance explained in a model is attributed to the measurement

method rather than the constructs the measures represent (Mackenzie *et al.*, 2005). In extreme cases, more than 50% of the explained variance can result from common method variance (Sharma *et al.*, 2009). Due to the fact that we used only one data source and gathered the data for the exogenous and endogenous constructs from the same participants, our study could have been affected by common method variance (Podsakoff *et al.*, 2003). To account for this problem, we followed the guidelines by Podsakoff *et al.* (2003) and used procedural remedies to reduce the probability that the common method variance would impact our results. We first assured anonymity to the participants by explicitly stating in the introduction of the questionnaire that all answers would be anonymous and that no relationship between any answers and participants would be established. Second, the introduction also stated that there were no right or wrong answers, emphasizing that we were interested in the participants' honest opinions. Third, we provided verbal labels for the extreme points and midpoints of the scales.
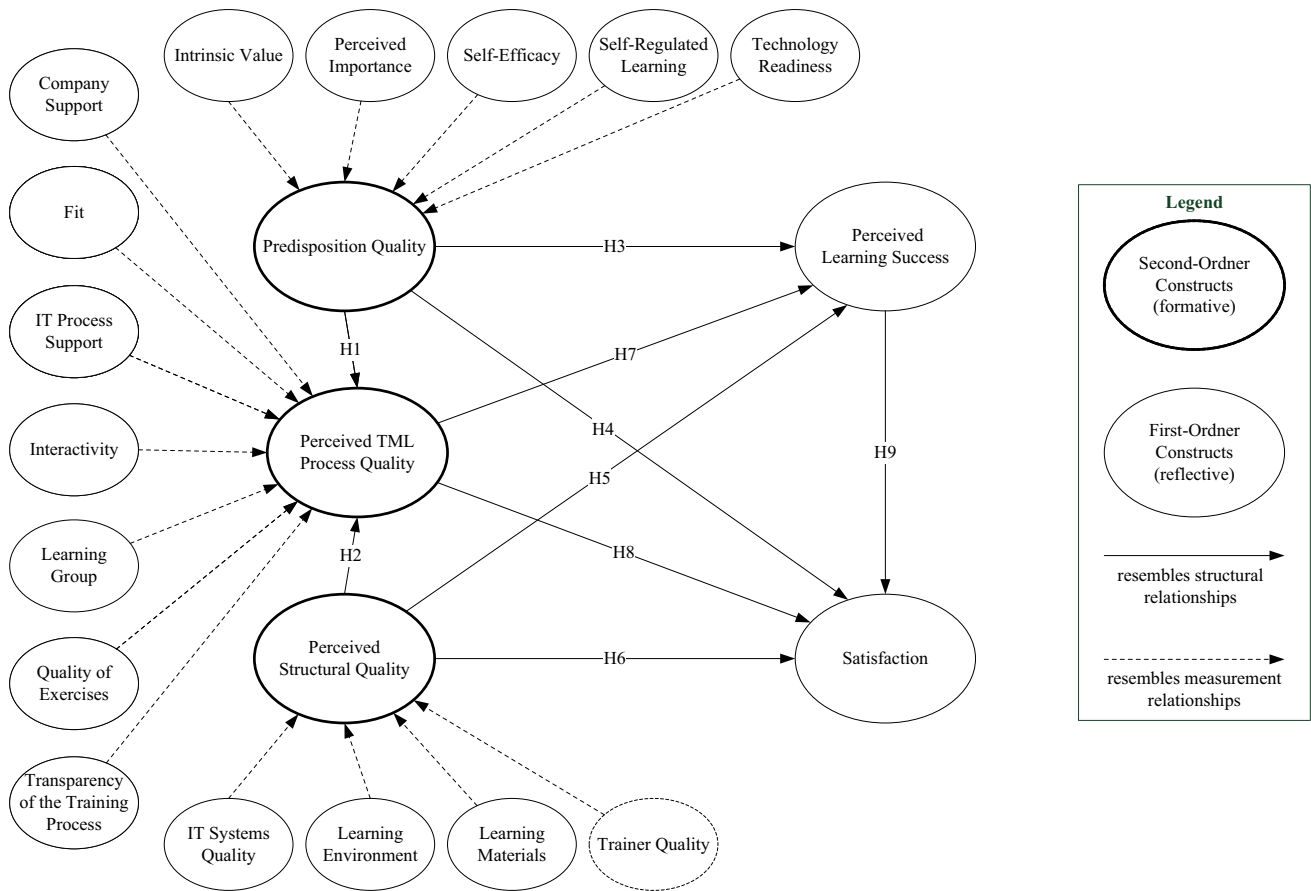
Figure 3 Research Model including the First-Order Constructs of the Formative Second-Order Constructs.

Fourth, we used two questionnaires that were completed at different points during each course to separate some of the exogenous and endogenous constructs. Furthermore, in the second questionnaire, we used a cover story to create the perception that the exogenous and endogenous constructs were not connected. Regarding the statistical remedies, we followed Sattler *et al.* (2010) and only conducted Harman's single-factor test but no other techniques, since all existing methods have shown to lack effectiveness in detecting common method variance. The results of the test show that the single factor does not account for more than half of the variance (in detail, it accounts for 35%), and thus, it is unlikely that our results are notably influenced by common method variance. The complete research model of our main study including the first-order constructs of the formative second-order constructs is illustrated in Figure 3.

Results

Measurement models

Due to the fact that we used reflective and formative measurement models and that both need to be evaluated using different quality criteria (Chin, 1998), we separately assessed the quality of the reflective and formative measurement models. Beginning with the evaluation of the reflective measurement models, we focus on the results for the lowest

indicator loading, the composite reliability (ρ_c), the average variance extracted (AVE), and the heterotrait–monotrait ratio (HTMT, Henseler *et al.*, 2015).

The results presented in Table 3 show that all loadings are higher than 0.76 (should be above 0.707). Additionally, the composite reliability for all constructs is higher than 0.80 (should be above 0.707), and the AVE for all constructs is higher than 0.62 (should be above 0.5). Furthermore, the HTMT ratio is well below 0.85, indicating the presence of discriminant validity (Henseler *et al.*, 2015).

After having shown that the reflective measurement models fulfill the desired quality criteria, we now focus on the evaluation of the formative measurement models. For this evaluation, we rely on the six guidelines for evaluating formative measurement models presented by Cenfetelli and Bassellier (2009). A summary of the key indicators is presented in Table 4 (for an in-depth analysis of the formative measurement models, please see "Appendix 2").

The results presented in Table 4 show that the formative measurement models fulfill the guidelines by Cenfetelli and Bassellier (2009). We observed problematic results for two indicators, since *Technology Readiness,* and *Learning Environment* show: (a) negative, (b) nonsignificant weights, and (c) low loadings. However, we followed the recommendation of Cenfetelli and Bassellier (2009) not to drop these indicators since their inclusion is well grounded in theory on TML. However, if subsequent studies observe similar issues with these indicators, their inclusion should be questioned.

Table 3 Cross-loadings and composite reliability for the reflective measurement models

	Perceived Learning Success	Satisfaction
Lowest indicator loading	0.761	0.793
ρ_c	0.871	0.808
AVE	0.628	0.678
HTMT	0.664	

In summary, the evaluation of our reflective and formative measurement models shows that they fulfill the desired quality criteria. Thus, we can now confidently move on to the evaluation of the structural model.

Structural model

Regarding the evaluation of the structural model, we follow the guidelines by Hair *et al.* (2014). First, we assess possible multicollinearity issues for endogenous constructs with two or more predictors (see Table 5). Since the highest variance inflation factor (VIF) value (3.555) is below the limit of 5 (Hair *et al.*, 2014), multicollinearity among the predictors of the endogenous constructs is not an issue in this study.

Figure 4 summarizes the results of the structural model relationships, the R^2 of the endogenous constructs, and the Q^2 of the reflectively measured endogenous constructs.

The Q^2 scores above 0 show that the structural model has predictive relevance for both TML outcomes: perceived learning success and satisfaction (Hair *et al.*, 2014). The R^2 values show that 56.3 and 67.8%, respectively, of the variance in perceived learning success and satisfaction can be explained, which is a good result. Regarding perceived TML process quality, 61.6% of the variance is explained by the two TML input factors. Furthermore, we found support for six of our nine hypotheses.

Due to the fact that significance alone is not an indicator of importance (Ringle *et al.*, 2012), we subsequently assessed the effect size f^2 of each relationship (Table 6). Using this measure, we can grasp the impact of omitting one predictor of an endogenous construct in terms of the change in the R^2 value of the construct. Additionally, Hair *et al.* (2014) recommend assessing the q^2 effect size of each relationship to compare the predictive relevance of the individual relationships (Table 6). Values of 0.02, 0.15, and 0.35 resemble a small, medium, or large f^2 or q^2 effect size,[2] respectively.

The results presented in Table 6 show that we found at least small f^2 effects for all significant relationships. Furthermore, we observed the largest f^2 effects for the relationships between perceived structural quality and perceived TML process quality (large), and between perceived TML process quality and perceived learning success as well as satisfaction (both medium). Regarding the q^2 effects, we found small positive effects between predisposition quality and perceived learning success, perceived TML process quality and perceived learning success, perceived TML process quality and satisfaction, perceived structural quality and satisfaction, and perceived learning success and satisfaction.

After presenting the path coefficients, significances, as well as the f^2 and q^2 effect sizes, we need to keep in mind that the path coefficients resemble only the direct effects between two constructs. Nevertheless, we also need to take the indirect effects into account for answering our research question regarding the impact of the different factors in a comprehensive evaluation of TML quality. Considering our construct predisposition quality, for example, this construct has a significant direct effect on perceived TML process quality but no significant direct effect on satisfaction. However, it would be incorrect to conclude that predisposition quality has no effects on satisfaction without investigating the indirect effects via perceived TML process quality. Table 7 summarizes the results regarding the total effects (direct + indirect effects).

Table 4 VIF, factor weights, *p* value, and factor loadings for the indicators of the formative measurement models

Construct	Indicator	VIF	Factor Weights	p value	Factor Loadings
Predisposition Quality	Perceived Importance	1.084	0.784	<0.001	
	Intrinsic Value	1.247	0.221	n.s.	0.498
	Self-Efficacy	1.670	0.196	n.s.	0.451
	Self-Regulated Learning	1.228	0.233	<0.05	
	Technology Readiness	1.555	−0.081	n.s.	0.228
Perceived TML Process Quality	Fit	1.729	0.420	<0.001	
	IT Process Support	Indicator dropped*			
	Interactivity	1.236	0.098	n.s.	0.489
	Learning Group	Indicator dropped*			
	Quality of Exercises	1.851	0.230	<0.01	
	Transparency	1.622	0.307	<0.001	
	Company Support	1.238	0.301	<0.001	
Perceived Structural Quality	IT Systems Quality	1.753	−0.090	n.s.	0.445
	Learning Environment	1.372	−0.044	n.s.	0.218
	Learning Materials	1.440	0.789	<0.001	
	Trainer Quality	1.199	0.467	<0.001	

*We dropped this indicator, since we observed a nonsignificant weights and a low loading. Furthermore, the indicator was developed in our focus groups and thus is not grounded in theory.

Table 5 Multicollinearity among the Predictors of Endogenous Constructs

Predictors of Perceived TML Process Quality		Predictors of Perceived Learning Success		Predictors of Satisfaction	
Construct	VIF	Construct	VIF	Construct	VIF
Predisposition Quality	1.428	Predisposition Quality	1.631	Predisposition Quality	1.721
Structural Quality	1.428	Structural Quality	2.354	Structural Quality	2.358
		TML Process Quality	2.602	TML Process Quality	3.555
				Learning Success	2.287
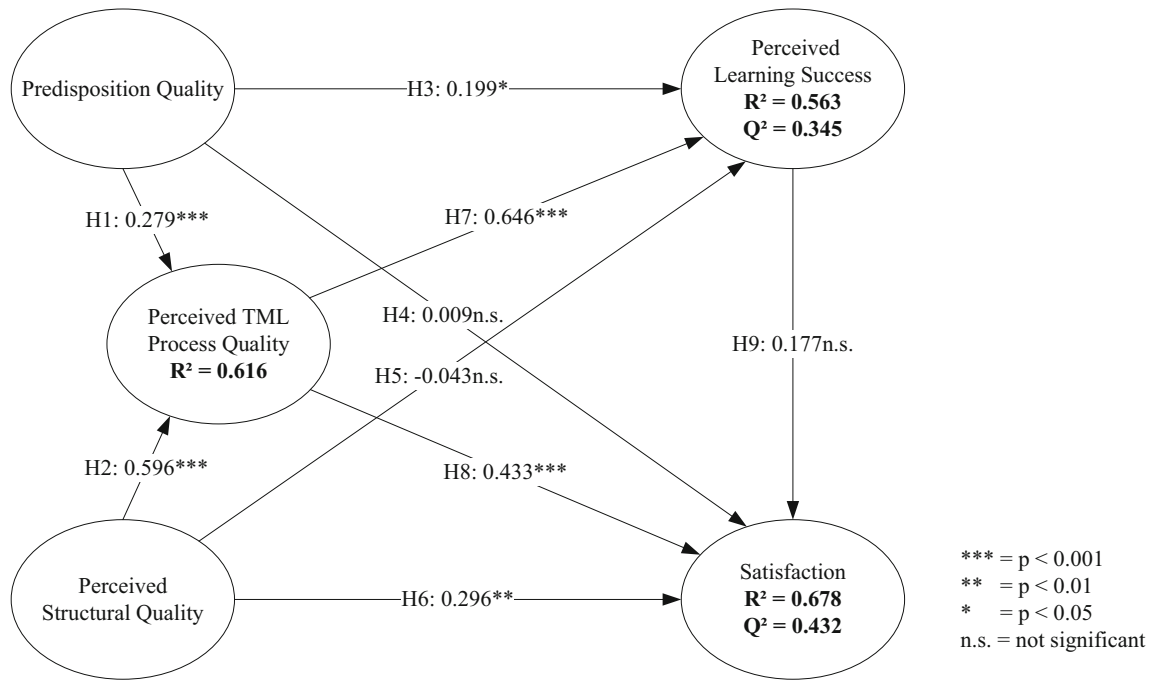
Figure 4 Evaluated Research Model.

Table 6 f^2 and q^2 effect sizes for the structural model

Relationship	f^2 value	f^2 effect Size	q^2 value	q^2 effect size
Predisposition Quality → Perceived TML Process Quality	0.135	Small		
Perceived Structural Quality → Perceived TML Process Quality	0.651	Large		
Predisposition Quality → Perceived Learning Success	0.055	Small	0.020	Small
Perceived TML Process Quality → Perceived Learning Success	0.327	Medium	0.139	Small
Perceived Structural Quality → Perceived Learning Success	−0.014		−0.009	
Predisposition Quality → Satisfaction	0.009		−0.014	
Perceived TML Process Quality → Satisfaction	0.174	Medium	0.067	Small
Perceived Structural Quality → Satisfaction	0.115	Small	0.044	Small
Perceived Learning Success → Satisfaction	0.078	Small	0.033	Small

The results presented in Table 7 provide insights into the accumulated impact of the different factors on TML outcomes. Even though we did not observe direct effects of predisposition quality on satisfaction, as well as of perceived structural quality on perceived learning success, the total effects show that this factor matters in a holistic TML performance evaluation. In fact, perceived structural quality has the highest total effect on satisfaction, followed by perceived TML process quality, predisposition quality, and perceived learning success ($0.615 > 0.547 > 0.197 > 0.177$). Regarding perceived learning success, perceived TML process quality has the highest impact, followed by predisposition quality and perceived structural quality with an almost equal impact ($0.646 > 0.379 > 0.342$).

Discussion

Our study shows that the often neglected TML process has a major influence on both considered TML outcomes, in specific, satisfaction and perceived learning success. As seen in our model analysis, perceived TML process quality as the key construct of our research model is truly one of the major

Table 7 Total effects for the structural model

Relationship	Direct Effect	Indirect Effect	Total Effect	t value
Predisposition Quality → Perceived TML Process Quality	0.279		0.279***	4.417
Perceived Structural Quality → Perceived TML Process Quality	0.596		0.596***	10.171
Predisposition Quality → Perceived Learning Success	0.199	0.180	0.379***	3.726
Perceived TML Process Quality → Perceived Learning Success	0.646		0.646***	6.485
Perceived Structural Quality → Perceived Learning Success	−0.043	0.385	0.342***	3.816
Predisposition Quality → Satisfaction	0.009	0.188	0.197*	2.366
Perceived TML Process Quality → Satisfaction	0.433	0.114	0.547***	5.100
Perceived Structural Quality → Satisfaction	0.296	0.319	0.615***	8.627
Perceived Learning Success → Satisfaction	0.177		0.177n.s.	1.333

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; n.s.: not significant.

constructs when considering the comprehensive performance assessment of TML. The complete evaluated research model including the first-order and second-order constructs is found in Figure 5. Below, we discuss the theoretical and practical contributions our study results provide.

Theoretical implications

Our paper makes several contributions to the existing body of literature. When answering our first research questions, we identified constructs and categories to evaluate TML performance. We contribute to information systems research in general as well as TML research in particular by providing a TML-specific evaluation approach that covers the input, process, and outcome perspectives necessary to evaluate TML holistically as requested by several researchers (Gupta and Bostrom, 2009; van Dyke *et al.*, 1997). By using the theoretical foundation provided by Gupta and Bostrom (2009) that explicitly considers the TML process, we were able to identify and adapt 19 facets for five identified constructs by means of a literature review, a focus group workshop, the application of the Q-sort method, as well as an empirical pre-study and main study. Thus, by conducting a rigorous scale development approach with TML practitioners, we contribute to research by providing a set of scales that is suitable for a wide range of TML assessments and should engage the ongoing discussion about which learning methods or facets of the learning process are key contributors to learning outcomes in various application domains of TML.

By means of this study, we are the first to examine a holistic TML performance model in the course of vocational training. Thereby, we enrich the existing body of literature by adding and sharpening dimensions that have previously been the focus of TML research. While research on the TML process dimension in a university context has been conducted before (Bitzer *et al.*, 2013; Yi and Davis, 2003; Gupta and Bostrom, 2013), we enrich existing findings in the literature using data from vocational training participants. We examined the TML process dimension, identifying components and corresponding items to examine the effects of structural quality characteristics and learners' predispositions via the TML process dimension. More precisely, we were able to identify the most relevant facets of perceived TML process quality such as the overall fit (factor weight = 0.420, $p < 0.001$), transparency of the learning process (factor weight = 0.307, $p < 0.001$), the quality of exercises (factor weight = 0.230, $p < 0.01$), and company support (factor weight = 0.301, $p < 0.001$). In doing so, we

successfully extended existing with TML-specific evaluation approaches and thus offer insights into the critical drivers of TML quality, as demanded by other researchers (Sultan and Wong, 2010). By means of these components, we enable TML researchers and practitioners to identify more specific weaknesses within TML scenarios, which go beyond classical, high-level dimensions of TML quality evaluation. Additionally, our findings challenge the role of interactivity as a key facet of the TML process (Bitzer and Janson, 2014; Arbaugh and Benbunan-Fich, 2007; Lehmann and Söllner, 2014; Siau *et al.*, 2006), since our results show that interactivity (factor weight = 0.098, n.s.) is not a significant facet of perceived TML process quality. Rather, the overall fit of the provided TML scenario should be especially considered when designing the TML process. As pointed out by Yoo *et al.* (2002), it is not about just introducing technology for the purpose of learning, which may only result in overwhelmed learners, thus interfering with the learning process. Thus, our results provide further evidence that providers of TML should concentrate to deliver TML that is considered from an overall fit perspective. Therefore, our research highlights the need to deliver a high-quality TML process that is well designed and explicitly integrates technology with the learning process.

In addition, we observed high and significant impacts of both the quality of the learning material (factor weight = 0.789, $p < 0.001$) and the trainer quality (factor weight = 0.467, $p < 0.001$) on the perceived structural quality of TML. They were of higher relevance than components such as the learning environment (factor weight = −0.044, n.s.) and IT system quality (factor weight = −0.090, n.s.). This seems surprising, since many studies have suggested a high impact of IT system quality (Lin, 2007) and the learning environment (Ladhari, 2009). Nevertheless, IT system quality may be a minimum requirement nowadays, considered as mandatory and which does not significantly influence the learning process.

Considering the facets included in predisposition quality, our results showed that the perceived importance was the strongest facet of predisposition quality (factor weight = 0.784, $p < 0.001$), followed by the ability for self-regulated learning (factor weight = 0.233, $p < 0.05$). Surprisingly, intrinsic value (factor weight = 0.221, n.s.), self-efficacy (factor weight = 0.196, n.s.), and technology readiness (factor weight = −0.081, n.s.) have not proved to be significant in influencing predisposition quality, although previous research has highlighted the role of both facets for the predisposition quality of learners in TML (see, e.g., van der Rhee *et al.*, 2007; Pintrich and De Groot, 1990).
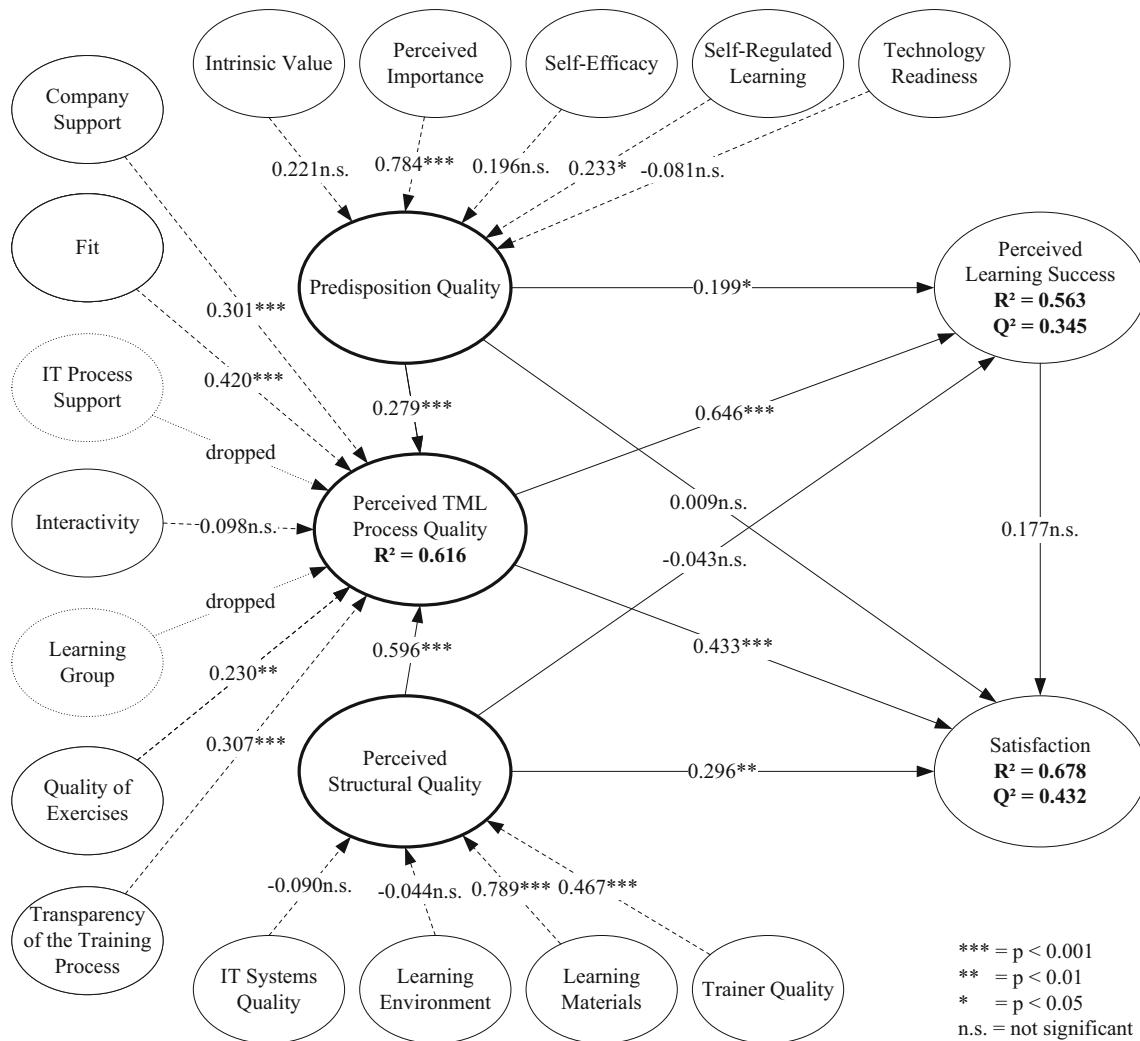
Figure 5 Evaluated Research Model including the First-Order Constructs of the Second-Order Constructs.

The second research question dealt with the relationship of the various constituting constructs of TML quality. We identified four dimensions resembled by the five constructs (1) perceived structural quality, (2) predisposition quality, (3) perceived TML process quality, (4) satisfaction, and (5) perceived learning success. In our research model, we showed that perceived TML process quality could help explain the inconclusive results observed in former research results (Gupta and Bostrom, 2009). We were able to show that both perceived structural quality (path coefficient = 0.596, $p < 0.001$) and predisposition quality (path coefficient = 0.279, $p < 0.001$) have a significant influence on perceived TML process quality.

Furthermore, we observed that perceived TML process quality is particularly important for determining the impact of perceived structural quality on TML outcomes. We did not observe a direct effect of perceived structural quality on perceived learning success (path coefficient = −0.043, n.s.), but observed a significant and high total impact via perceived TML process quality (total effect = 0.342, $p < 0.001$). Despite the fact that the existence of an indirect effect is especially important in this particular case, we observed significant indirect effects for both independent variables on one of the service outcomes (perceived structural quality on perceived learning success, and predisposition quality on satisfaction). Combining these indirect effects with the direct effects of TML process quality on learning success (0.646, $p < 0.001$) and satisfaction (0.433, $p < 0.001$), the results of our study empirically validate the central role of perceived TML process quality, as highlighted in previous conceptual TML research (Gupta and Bostrom, 2009) as well as empirical research (Janson *et al.*, 2017).

In sum, we provided a comprehensive theoretical model that corresponds with Gregor's (2006) theory of explanation and prediction, since it explains and predicts factor effects within TML. We were able to answer our research questions by presenting a TML evaluation approach based on a comprehensive model of five constructs, namely perceived structural quality, predisposition quality, perceived TML process quality, perceived learning success and satisfaction. We developed a holistic model that considers the specific requirements of a holistic TML evaluation. In addition, we were able to demonstrate relationships between the five constructs, providing insights into causal effects of the various constructs and delivering evidence of the importance of the TML processes. Furthermore, we provided a deeper

look into the single facets forming perceived structural quality, predisposition quality and perceived TML process quality.

Practical implications

Furthermore, our results help TML providers and trainers to evaluate the performance of TML more adequately, thereby considering comprehensive TML quality-related components. The effects of various treatments and antecedents, such as predisposition quality and perceived structural quality, can be observed and taken into account for redesigning TML. Moreover, the simultaneous evaluation of five related constructs empowers researchers and practitioners to derive and survey multidimensional measures, supporting an efficient and effective TML delivery. Furthermore, a continuous improvement process that derives improvement measures for the quality of the TML process while taking both perceived structural and predisposition quality into account is encouraged.

Considering our results, practitioners deploying vocational software trainings should heavily concentrate on supporting the TML process. In this context, scaffolds known from educational research serve as a design implication initially supporting the learners in their learning process (Gupta and Bostrom, 2009; Delen *et al.*, 2014; Janson and Thiel de Gafenco, 2015), for example, by providing learning paths, thus fostering TML quality by preventing learners from being overwhelmed by large amounts of learning material and allowing them to focus on the learning itself by means of initial support in the learning process.

Also, our results show that TML providers should aim at aligning their offerings and the needs of the participants to ensure a high fit. Thus, TML providers should consider how they integrate TML in the learning process of individuals by tailoring TML processes to the individual needs. Solutions that can be considered are, for example, the design of adaptive learning paths for better matching the expectations of learners in TML, thus ensuring a high-quality TML process.

Limitations and future research

We acknowledge several limitations to this study, which then underline a demand for future research. We investigated TML performance in a vocational software training. For this reason, the external validity of this study could be endangered (Bordens and Abbott, 2011), since the results of our empirical study might not be fully generalizable to other TML research contexts, for example, TML scenarios such as MOOCs in higher education. Hence, future research should see our research as a starting point and use our TML performance model in other research contexts to further evaluate our model and to gain additional insights regarding the structural relationships of the constituting TML dimensions in other research contexts.

The study examined the TML process with a single measurement and did therefore not account for the temporal development of the TML process (Gupta and Bostrom, 2009). Though we conducted a study with two measurement points, we did only capture a snapshot of the TML process and its determinants. By this means, we contribute to theory by identifying causal relationships and testing generalizable hypotheses between TML constructs and their influence on TML quality. Hence, the

necessity to conduct longitudinal studies regarding how the TML process evolves over time arises. Extending our variance approach used in this study, future research should conduct panel analyses to study the TML process and its influence on TML outcomes more thoroughly (Poole *et al.*, 2000). In addition, process approaches such as narratives would also contribute to a further understanding of how the TML process relates to the quality of TML. In contrast to our variance approach, a process approach should enable researchers to investigate a detailed reconstruction of individual TML processes and how specific events in the TML process influence TML quality (Poole *et al.*, 2000).

The measurement of TML quality components is connected to these issues. In this particular study, we relied on subjective measures, that is, TML participant perceptions. However, future TML research should consider objective measures. This particular limitation also applies to our measurement of learning success with the scales of Alavi (1994) and refers to an ongoing discussion regarding the suitability of self-reported or objective learning success data (Benbunan-Fich, 2010; Sitzmann *et al.*, 2010; Janson *et al.*, 2014). Hence, future research should also use objective measures for learning success as a dependent variable, such as cognitive and skill-based learning outcomes, which relate to how people are exhibiting changes in cognitive knowledge immediately after a training as well as how they perform at their workplace (Yi and Davis, 2003; Santhanam *et al.*, 2013, 2016). Therefore, we acknowledge that our study results concerning learning success may not directly refer to the actual evidence of learning. However, on the one hand, such meta-cognitive measures are as a proxy often intertwined with knowledge acquisition, and on the other hand, the model itself is also applicable to contexts where a measurement of cognitive knowledge acquisition is possible by simply replacing the latent construct with test scores of individual learners. Furthermore, literature on vocational trainings has highlighted the importance of participant perception, and satisfaction with the training, since these factor lead to loyalty with a specific training provider (Kuo and Ye, 2009). Since marketing literature has shown that satisfied and loyal customers are more likely to engage in positive word-of-mouth (Matos and Rossi, 2008), focusing on ensuring positive perceptions of the participants toward the training and the provider is a key success factor for vocational training providers, due to the positive influence on future revenues.

The last limitation and according area for future research involve the data collection and measurement. We used the same instrument to assess the dependent and independent latent variables among all participants of the study, allowing for possible common method variances, which should be addressed in future research. Nevertheless, procedural remedies were taken in this study to avoid ex ante biases, and Harman's single-factor test indicated that common method variance is not an issue in our study. However, since statistical tests for common method variances have their limitations (Chin *et al.*, 2012), we cannot rule out the existence of common method variance completely.

Conclusion

In this paper, we aimed to broaden the body of knowledge in the context of TML performance. For this purpose, we developed a model for the holistic evaluation of TML scenarios

with a special focus on vocational software trainings. Based on the literature, we derived components for the evaluation of TML performance, thereby, respectively, identifying and confirming the existence of new components that have not been considered thus far in the empirical literature, mainly for the TML process dimension. In this dimension, we could add to the body of literature by identifying components such as transparency of the training process, quality of exercises, fit, and company support, apart from interactivity, which is widely known as a driver of TML process quality. Thus, for the first time, the foundation for comprehensive TML performance evaluation has been built, including extensive information on TML process quality. Additionally, we showed a relationship between the five TML performance-related constructs, providing arguments that a TML performance evaluation should follow a holistic approach, accounting for the importance of the TML input, process, and outcome perspectives. With the consideration of the TML process quality dimension, a comprehensive evaluation of TML quality is feasible and possesses the potential to adjust existing shortcomings regarding the examination and evaluation of TML scenarios.

Acknowledgements

Notes
1 The insights presented in this subsection resemble the results of our literature review (step 1a of our research process). For methodological details on how the review was conducted, please see the section Research Method.
2 We used the blindfolding algorithm of SmartPLS to compute the q^2 effect sizes. This algorithm can only handle reflectively measured endogenous constructs. Thus, we could compute the q^2 effect sizes for predictors of only two of our endogenous constructs.

References
Alavi, M. (1994). Computer-Mediated Collaborative Learning: An Empirical Evaluation. *MIS Quarterly,* **18**(2), 159–174.

Alavi, M. and Leidner, D.E. (2001a). Research Commentary: Technology-Mediated Learning–A Call for Greater Depth and Breadth of Research. *Information Systems Research,* **12**(1), 1–10.

Alavi, M. and Leidner, D.E. (2001b). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues, *MIS Quarterly*: 107–136.

Alavi, M., Marakas, G.M. and Yoo, Y. (2002). A comparative study of distributed learning environments on learning outcomes. *Information Systems Research,* **13**(4), 404–415.

Alavi, M., Wheeler, B.C. and Valacich, J.S. (1995). Using IT to Reengineer Business Education: An Exploratory Investigation of Collaborative Telelearning. *MIS Quarterly,* **19**(3), 293–312.

Alavi, M., Yoo, Y. and Vogel, D.R. (1997). Using Information Technology to Add Value to Management Education. *The Academy of Management Journal,* **40**(6), 1310–1333.

Alonso, F., Manrique, D., Martinez, L. and Vines, J.M. (2011). How Blended Learning Reduces Underachievement in Higher Education: An Experience in Teaching Computer Sciences. *IEEE Transactions on Education,* **54**(3), 471–478.

Aragon-Sanchez, A., Barba-Aragón, I. and Sanz-Valle, R. (2003). Effects of training on business results1. *The International Journal of Human Resource Management,* **14**(6), 956–980.

Arbaugh, J.B. (2000). How Classroom Environment and Student Engagement Affect Learning in Internet-based MBA Courses. *Business Communication Quarterly,* **63**(4), 9–26.

Arbaugh, J.B. (2001). How Instructor Immediacy Behaviors Affect Student Satisfaction and Learning in Web-Based Courses. *Business Communication Quarterly,* **64**(4), 42–54.

Arbaugh, J.B. (2013). Does Academic Discipline Moderate CoI-Course Outcomes Relationships in Online MBA Courses? *The Internet and Higher Education,* **17**, 16–28.

Arbaugh, J.B. and Benbunan-Fich, R. (2007). The importance of participant interaction in online environments. *Decision Support Systems,* **43**(3), 853–865.

Armstrong, S.J., Allinson, C.W. and Hayes, J. (2004). The effects of cognitive style on research supervision: A study of student-supervisor dyads in management education. *Academy of Management Learning and Education,* **3**(1), 41–63.

Becker, J.-M., Klein, K. and Wetzels, M. (2012). Hierarchical Latent Variable Models in PLS-SEM: Guidelines for Using Reflective-Formative Type Models. *Long Range Planning,* **45**(5–6), 359–394.

Benbunan-Fich, R. (2010). Is Self-Reported Learning a Proxy Metric for Learning? Perspectives From the Information Systems Literature. *Academy of Management Learning and Education,* **9**(2), 321–328.

Benbunan-Fich, R. and Arbaugh, J.B. (2006). Separating the Effects of Knowledge Construction and Group Collaboration in Learning Outcomes of Web-based Courses. *Information and Management,* **43**(6), 778–793.

Biggs, J. (1993). What do inventories of students' learning processes really measure?: A theoretical review and clarification. *British Journal of Educational Psychology,* **63**(1), 3–19.

Biggs, J., Kember, D. and Leung, D.Y. (2001). The revised two-factor Study Process Questionnaire: R-SPQ-2F. *British Journal of Educational Psychology,* **71**(1), 133–149.

Bitzer, P. and Janson, A. (2014). Towards a Holistic Understanding of Technology-Mediated Learning Services – a State-of-the-Art Analysis, *ECIS 2014 Proceedings*.

Bitzer, P., Söllner, M. and Leimeister, J.M. (2013). Evaluating the Quality of Technology-Mediated Learning Services, *ICIS 2013 Proceedings*.

Bordens, K.S. and Abbott, B.B. (2011). *Research design and methods: A process approach* (8th ed.). New York: McGraw-Hill.

Brower, H.H. (2003). On emulating classroom discussion in a distance-delivered OBHR course: Creating an on-line learning community. *Academy of Management Learning and Education,* **2**(1), 22–36.

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

Chin, W. (1998). The partial least squares approach for structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). Mahwah, N.J: Lawrence Erlbaum.

Chin, W.W., Thatcher, J.B. and Wright, R.T. (2012). Assessing common method bias: Problems with the ULMC technique. *MIS Quarterly,* **36**(3), 1003–1019.

Choi, D.H., Kim, J. and Kim, S.H. (2007). ERP Training with a Web-Based Electronic Learning System: The Flow Theory Perspective. *International Journal of Human-Computer Studies,* **65**(3), 223–243.

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

Chwelos, P., Benbasat, I. and Dexter, A.S. (2001). Research Report: Empirical Test of an EDI AdoptionModel. *Information Systems Research,* **12**(3), 304.

Cole, M.S., Feild, H.S. and Harris, S.G. (2004). Student learning motivation and psychological hardiness: Interactive effects on students' reactions to a management class. *Academy of Management Learning and Education,* **3**(1), 64–85.

Colquitt, J.A., LePine, J.A. and Noe, R.A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology,* **85**(5), 678–707.

de Matos, C.A. and Rossi, C.A.V. (2008). Word-of-mouth communications in marketing: A meta-analytic review of the antecedents and moderators. *Journal of the Academy of Marketing Science*, **36**(4), 578–596.

Deci, E.L. and Ryan, R.M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, **11**(4), 227–268.

Delen, E., Liew, J. and Willson, V. (2014). Effects of interactivity and instructional scaffolding on learning: Self-regulation in online video-based environments. *Computers and Education*, **78**, 312–320.

DeLone, W.H. and McLean, E.R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, **3**(1), 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), 9–30.

Dijkstra, T. and Henseler, J. (2015). Consistent Partial Least Squares Path Modeling. *Management Information Systems Quarterly*, **39**(2), 297–316.

Eom, S.B. (2011). Relationships Among E-Learning Systems and E-Learning Outcomes: A Path Analysis Model. *Human Systems Management*, **30**(4), 229–241.

Eom, S. (2015). Effects of Self-Efficacy and Self-regulated Learning on LMS User Satisfaction and LMS Effectiveness, *AMCIS 2015 Proceedings*.

Eom, S.B., Wen, H.J. and Ashill, N. (2006). The Determinants of Students' Perceived Learning Outcomes and Satisfaction in University Online Education: An Empirical Investigation*. *Decision Sciences Journal of Innovative Education*, **4**(2), 215–235.

Evans, C. and Gibbons, N.J. (2007). The interactivity effect in multimedia learning. *Computers and Education*, **49**(4), 1147–1160.

Frey, J.H. and Fontana, A. (1991). The group interview in social research. *The Social Science Journal*, **28**(2), 175–187.

Garrison, D.R. and Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education*, **7**(2), 95–105.

Gregor, S. (2006). The nature of theory in information systems. *MIS Quarterly*, **30**(3), 611–642.

Gupta, S. and Bostrom, R. (2009). Technology-Mediated Learning: A Comprehensive Theoretical Model. *Journal of the Association for Information Systems*, **10**(9), 686–714.

Gupta, S. and Bostrom, R. (2013). An Investigation of the Appropriation of Technology-Mediated Training Methods Incorporating Enactive and Collaborative Learning. *Information Systems Research*, **24**(2), 454–469.

Gupta, S., Bostrom, R.P. and Huber, M. (2010). End-user Training Methods: What We Know, Need to Know,. *SIGMIS Database*, **41**(4), 9–39.

Hair, J.F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M. (2014). *A primer on partial least squares structural equations modeling (PLS-SEM)*. Los Angeles: Sage.

Hannafin, M.J., Kim, M.C. and Kim, H. (2004). Reconciling research, theory, and practice in web-based teaching and learning: The case for grounded design. *Journal of Computing in Higher Education*, **15**(2), 3–20.

Hansen, D.E. (2008). Knowledge Transfer in Online Learning Environments. *Journal of Marketing Education*, **30**(2), 93–105.

Hattie, J. and Yates, G.C.R. (2014). *Visible learning and the science of how we learn*. New York: Routledge.

Henseler, J., Ringle, C.M. and Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, **43**(1), 115–135.

Hiltz, S.R., Coppola, N., Rotter, N., Toroff, M. and Benbunan-Fich, R. (2000). Measuring the importance of collaborative learning for the effectiveness of ALN: A multi-measure. *Online Education: Learning Effectiveness and Faculty Satisfaction*, **1**, 101–119.

Jacobs, S.C., Huprich, S.K., Grus, C.L., Cage, E.A., Elman, N. S., Forrest, L., et al. (2011). Trainees with professional competency problems: Preparing trainers for difficult but necessary conversations. *Training and Education in Professional Psychology*, **5**(3), 175.

Janson, A., Söllner, M., Bitzer, P. and Leimeister, J.M. (2014). Examining the Effect of Different Measurements of Learning Success in Technology-mediated Learning Research, *ICIS 2014 Proceedings*.

Janson, A., Söllner, M. and Leimeister, J.M. (2017). Individual Appropriation of Learning Management Systems: Antecedents and Consequences, *AIS Transactions on Human-Computer Interaction* (forthcoming).

Janson, A. and Thiel de Gafenco, M. (2015). Engaging the Appropriation of Technology-mediated Learning Services – A Theory-driven Design Approach, *ECIS 2015 Proceedings*.

Johnson, R.D., Gueutal, H. and Falbe, C.M. (2009). Technology. Trainees, Metacognitive Activity and E-Learning Effectiveness *Journal of Managerial Psychology*, **24**(6), 545–566.

Kim, J., Kwon, Y. and Cho, D. (2011). Investigating Factors that Influence Social Presence and Learning Outcomes in Distance Higher Education. *Computers and Education*, **57**(2), 1512–1520.

Kirkpatrick, D. and Kirkpatrick, J. (2005). Transferring learning to behavior: Using the four levels to improve performance: Berrett-Koehler Publishers.

Klein, H.J., Noe, R.A. and Wang, C. (2006). Motivation to Learn and Course Outcomes: The Impact of Delivery Mode, Learning Goal Orientation, and Perceived Barriers and Enablers. *Personnel Psychology*, **59**(3), 665–702.

Kolfschoten, G.L., Briggs, R.O., de Vreede, G.-J., Jacobs, P.H.M. and Appelman, J.H. (2006). A conceptual foundation of the thinkLet concept for Collaboration Engineering. *International Journal of Human-Computer Studies*, **64**(7), 611–621.

Kuo, Y.-K. and Ye, K.-D. (2009). The causal relationship between service quality, corporate image and adults' learning satisfaction and loyalty: A study of professional training programmes in a Taiwanese vocational institute. *Total Quality Management and Business Excellence*, **20**(7), 749–762.

Ladhari, R. (2009). A review of twenty years of SERVQUAL research. *International Journal of Quality and Service Sciences*, **1**(2), 172–198.

Landis, J.R. and Koch, G.G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, **33**(1), 159.

Lehmann, K. and Söllner, M. (2014). Theory-driven design of a mobile-learning application to support different interaction types in large-scale lectures, *ECIS 2014 Proceedings*.

Lehmann, K., Söllner, M. and Leimeister, J.M. (2016). Design and Evaluation of an IT-based Peer Assessment to Increase Learner Performance in Large-Scale Lectures, *ICIS 2016 Proceedings*.

Leidner, D. and Jarvenpaa, S. (1995). The Use of Information Technology to Enhance Management School Education: A Theoretical View, *Management Information Systems Quarterly* **19**(3).

Lim, H., Lee, S.-G. and Nam, K. (2007). Validating E-Learning Factors Affecting Training Effectiveness. *International Journal of Information Management*, **27**(1), 22–35.

Lin, H.-F. (2007). Measuring Online Learning Systems Success: Applying the Updated DeLone and McLean Model. *CyberPsychology and Behavior*, **10**(6), 817–820.

Linnenbrink, E.A. and Pintrich, P.R. (2002). Motivation as an enabler for academic success. *School Psychology Review*, **31**(3), 313.

López-Pérez, M.V., Pérez-López, M.C. and Rodríguez-Ariza, L. (2011). Blended learning in higher education: Students' perceptions and their relation to outcomes. *Computers and Education*, **56**(3), 818–826.

Mackenzie, S.B., Podsakoff, P.M. and Jarvis, C.B. (2005). The Problem of Measurement Model Misspecification in Behavioral and Organizational Research and Some Recommended Solutions. *Journal of Applied Psychology*, **90**(4), 710–730.

Miles, M.B. and Huberman, A.M. (1994). Qualitative data analysis: An expanded sourcebook: Sage.

Moore, M.G. (1989). Editorial: Three types of interaction. *American Journal of Distance Education*, **3**(2), 1–7.

Nahm, A.Y., Rao, S.S., Solis-Galvan, L.E. and Ragu-Nathan, T.S. (2002). The Q-sort method: assessing reliability and construct validity of questionnaire items at a pre-testing stage. *Journal of Modern Applied Statistical Methods*, **1**(1), 15.

Ozkan, S. and Koseler, R. (2009). Multi-dimensional Students' Evaluation of E-Learning Systems in the Higher Education Context: An Empirical Investigation. *Computers and Education*, **53**(4), 1285–1296.

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

Petter, S., DeLone, W. and McLean, E. (2008). Measuring information systems success: models, dimensions, measures, and interrelationships. *European Journal of Information Systems*, **17**(3), 236–263.

Piccoli, G., Ahmad, R. and Ives, B. (2001). Web-Based Virtual Learning Environments: A Research Framework and a Preliminary Assessment of Effectiveness in Basic IT Skills Training, *Management Information Systems Quarterly* **25**(4).

Pintrich, P.R. and De Groot, E.V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, **82**(1), 33.

Podsakoff, P., MacKenzie, S., Lee, J. and Podsakoff, N. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*, **88**(5), 879–903.

Poole, M.S., de Ven, Van, Andrew, H., Dooley, K. and Holmes, M.E. (2000). *Organizational change and innovation processes: Theory and methods for research.* Oxford, New York: Oxford Univ. Press.

Rasch, T. and Schnotz, W. (2009). Interactive and Non-Interactive Pictures in Multimedia Learning Environments: Effects on Learning Outcomes and Learning Efficiency. *Learning and Instruction,* **19**(5), 411–422.

Reber, R.A. and Wallin, J.A. (1984). The effects of training, goal setting, and knowledge of results on safe behavior: A component analysis. *Academy of Management Journal,* **27**(3), 544–560.

Ringle, C.M., Sarstedt, M. and Straub, D.W. (2012). A Critical Look at the Use of PLS-SEM in MIS Quarterly, *MIS Quarterly* **36**(1): iii–xiv; s3–s8 (supplement).

Ringle, C.M., Wende, S. and Becker, J.-M. (2015). SmartPLS [WWW document] www.smartpls.de.

Rubin, D.B. (1987). *Multiple imputation for nonresponse in surveys.* New York: Wiley.

Santhanam, R., Liu, D. and Shen, W.-C.M. (2016). Research Note – Gamification of Technology-Mediated Training: Not All Competitions Are the Same. *Information Systems Research,* **27**(2), 453–465.

Santhanam, R., Sasidharan, S. and Webster, J. (2008). Using Self-Regulatory Learning to Enhance E-Learning-Based Information Technology Training. *Information Systems Research,* **19**(1), 26–47.

Santhanam, R., Yi, M., Sasidharan, S. and Park, S.-H. (2013). Toward an Integrative Understanding of Information Technology Training Research across Information Systems and Human-Computer Interaction: A Comprehensive Review. *AIS Transactions on Human-Computer Interaction,* **5**(3), 134–156.

Sattler, H., Völckner, F., Riediger, C. and Ringle, C.M. (2010). The Impact of Brand Extension Success Factors on Brand Extension Price Premium. *International Journal of Research in Marketing,* **27**(4), 319–328.

Schank, R. (2005). *Lessons in Learning, e-Learning, and Training. Perspectives and Guidance for the Enlightened Trainer.*, Pfeiffer: John Wiley.

Seyda, S. and Werner, D. (2012). IW-Weiterbildungserhebung 2011: Gestiegenes Weiterbildungsvolumen bei konstanten Kosten. *IW-Trends–Vierteljahresschrift zur empirischen Wirtschaftsforschung,* **39**(1), 37–54.

Sharma, R., Yetton, P. and Crawford, J. (2009). Estimating the Effect of Common Method Variance: The Method-Method Pair Technique with an Illustration from TAM Research. *MIS Quarterly,* **33**(3), 473.

Siau, K., Sheng, Hong and Nah, F.F.-H. (2006). Use of a classroom response system to enhance classroom interactivity. *Education, IEEE Transactions on,* **49**(3), 398–403.

Sims, R. (2003). Promises of Interactivity: Aligning Learner Perceptions and Expectations With Strategies for Flexible and Online Learning. *Distance Education,* **24**(1), 87–103.

Sitzmann, T., Ely, K., Brown, K.G. and Bauer, K.N. (2010). Self-Assessment of Knowledge: A Cognitive Learning or Affective Measure? *Academy of Management Learning and Education,* **9**(2), 169–191.

Smith, S.M. and Woody, P.C. (2000). Interactive Effect of Multimedia Instruction and Learning Styles. *Teaching of Psychology,* **27**(3), 220–223.

Söllner, M., Hoffmann, A., Hoffmann, H., Wacker, A. and Leimeister, J.M. (2012). Understanding the Formation of Trust in IT Artifacts, *ICIS 2016 Proceedings.*

Söllner, M., Hoffmann, A. and Leimeister, J.M. (2016). Why Different Trust Relationships Matter for Information Systems Users. *European Journal of Information Systems,* **25**(3), 274–287.

Sousa, N., Alves, G.R. and Gericota, M.G. (2010). An Integrated Reusable Remote Laboratory to Complement Electronics Teaching. *IEEE Transactions on Learning Technologies,* **3**(3), 265–271.

Stevens, J.P. (1992). *Applied Multivariate Statistics for the Social Sciences,* 2nd ed., Hillsdale: Lawrence Erlbaum Associates.

Straub, D.W. (1989). Validating Instruments in MIS Research. *MIS Quarterly,* **13**(2), 147–169.

Sultan, P. and Wong, H.Y. (2010). Service quality in higher education – a review and research agenda. *International Journal of Quality and Service Sciences,* **2**(2), 259–272.

Sun, P.-C., Tsai, R.J., Finger, G., Chen, Y.-Y. and Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. *Computers and Education,* **50**(4), 1183–1202.

Sweller, J. and Chandler, P. (1991). Evidence for Cognitive Load Theory. *Cognition and Instruction,* **8**(4), 351–362.

Tannenbaum, S.I., Mathieu, J.E., Salas, E. and Cannon-Bowers, J.A. (1991). Meeting trainees' expectations: The influence of training fulfillment on the development of commitment, self-efficacy, and motivation. *Journal of Applied Psychology,* **76**(6), 759.

Thurmond, V. and Wambach, K. (2004). Understanding interactions in distance education: A review of the literature. *International Journal of Instructional Technology and Distance Learning,* **1**(1), 9–26.

Tse, D.K. and Wilton, P.C. (1988). Models of Consumer Satisfaction Formation: An Extension. *Journal of Marketing Research,* **25**(2), 204–212.

van der Rhee, B., Verma, R., Plaschka, G.R. and Kickul, J.R. (2007). Technology readiness, learning goals, and eLearning: Searching for synergy. *Decision Sciences Journal of Innovative Education,* **5**(1), 127–149.

van Dyke, T.P., Kappelman, L.A. and Prybutok, V.R. (1997). Measuring information systems service quality: concerns on the use of the SERVQUAL questionnaire, *MIS Quarterly,* 195–208.

Vollmar, M. (2013). Berufliche Weiterbildung in Unternehmen 2010, Methodik und erste Ergebnisse. *Wirtschaft und Statistik, H,* **4,** 276–287.

Wang, Y.-S. (2003). Assessment of learner satisfaction with asynchronous electronic learning systems. *Information and Management,* **41**(1), 75–86.

Wang, W. and Benbasat, I. (2005). Trust in and adoption of online recommendation agents. *Journal of the Association for Information Systems,* **6**(3), 72–101.

Wang, Y.-S., Wang, H.-Y. and Shee, D.Y. (2007). Measuring E-Learning Systems Success in an Organizational Context: Scale Development and Validation. *Computers in Human Behavior,* **23**(4), 1792–1808.

Wegener, R., Menschner, P. and Leimeister, J. (2012). Design and Evaluation of a Didactical Service Blueprinting Method For Large Scale Lectures, *ICIS 2012 Proceedings.*

Wold, H.O.A. (1982). Soft Modeling: The Basic Design and Some Extensions. In K. G. Jöreskog and H. O. A. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction* (pp. 1–54). Amsterdam: North-Holland.

Yi, M.Y. and Davis, F.D. (2003). Developing and Validating an Observational Learning Model of Computer Software Training and Skill Acquisition: Information Systems Research. *Information Systems Research,* **14**(2), 146–169.

Yoo, Y., Kanawattanachai, P. and Citurs, A. (2002). Forging into the Wired Wilderness: A Case Study of a Technology-Mediated Distributed Discussion-Based Class. *Journal of Management Education,* **26**(2), 139–163.

About the Authors

Matthias Söllner is Assistant Professor at the Institute of Information Management (IWI-HSG) at University of St.Gallen, Switzerland, as well as Postdoctoral Researcher at the Information Systems department at the Research Center for IS Design (ITeG) at University of Kassel, Germany. His research focuses on understanding and designing successful systems in domains such as higher education and vocational training.

Philipp Bitzer (www.smartransfer.eu) received a Ph.D. in Information Systems from University of Kassel in 2014. His research is on quality and productivity aspects of IT-supported vocational trainings. Currently he focuses on transferring his research findings into practice as co-founder and CEO of a Startup called smarTransfer.

Andreas Janson is a Researcher and Ph.D. candidate at the Informations Systems department and Research Center for IS Design (ITeG) at University of Kassel, Germany. His research focuses on the appropriation and theory-driven design of technology-mediated learning in various application domains.

Jan Marco Leimeister is Director and Full Professor at the Institute of Information Management (IWI-HSG) at University of St.Gallen, Switzerland, and is also the Chair for Information Systems at the Research Center for IS Design at University of Kassel, Germany. His research focuses on Service Engineering, Collaboration Engineering, IT Innovations and Crowdsourcing.

Appendix 1: Measurement instrument used in the main study

Subdimension	Item	Statement (loading)	Source	Mean	SD
Perceived Structural Quality (reflective first-order, formative second-order construct)					
Trainer Quality $\rho_c = 0.902$ AVE = 0.607	TQ1	The instructor can answer in-depth questions easily (0.784)	Developed during the first step of our research process	4.27	0.81
	TQ2	The instructor can create a beneficial learning atmosphere (0.827)	Developed during the first step of our research process	4.17	0.80
	TQ3	The instructor can convey complex content comprehensibly (0.840)	Developed during the first step of our research process	4.10	0.90
	TQ4	The instructor knows how to build a relationship with the participants (0.686)	Parasuraman *et al.* (1985)	4.19	0.87
	TQ5	The instructor can assess the learning progress of the group very well (0.778)	Developed during the first step of our research process	3.93	0.85
	TQ6	The instructor is motivated (0.749)	Parasuraman *et al.* (1985)	4.46	0.75
	TQ7*	The instructor is always ready to help	Parasuraman *et al.* (1985)	4.75	0.60
Learning Environment $\rho_c = 0.928$ AVE = 0.812	LE1	The course room is modern (0.903)	Parasuraman *et al.* (1985)	4.38	0.81
	LE2	The equipment is optically appealing (0.898)	Parasuraman *et al.* (1985)	4.05	1.06
	LE3	The equipment of the course room appears to be of high quality (0.902)	Parasuraman *et al.* (1985)	3.87	1.14
IT Systems Quality $\rho_c = 0.903$ AVE = 0.651	SQ1	The provided IT systems are always available (0.737)	DeLone and McLean (2003)	4.07	1.09
	SQ2	The provided IT systems are reliable (0.837)	DeLone and McLean (2003)	4.05	0.99
	SQ3	The provided IT systems have an attractive user interface (0.830)	DeLone and McLean (2003)	3.90	0.97
	SQ4	The provided IT systems worked without delay (0.841)	DeLone and McLean (2003)	3.79	1.14
	SQ5	The provided IT systems are easy to use (0.785)	DeLone and McLean (2003)	4.12	0.86
Learning Materials $\rho_c = 0.867$ AVE = 0.567	LM1	The provided learning content was easy to understand (0.665)	DeLone and McLean (2003)	3.34	0.93
	LM2	The provided learning content was relevant (0.663)	DeLone and McLean (2003)	3.94	0.83
	LM3	The provided learning contents are well structured (0.845)	DeLone and McLean (2003)	3.73	0.89
	LM4	The provided learning contents are complete (0.752)	DeLone and McLean (2003)	3.99	0.93
	LM5	The provided learning contents are presented appealingly (0.822)	DeLone and McLean (2003)	3.53	0.99
Predisposition Quality (reflective first-order, formative second-order construct)					
Self-Efficacy $\rho_c = 0.849$ AVE = 0.654	**SE1**	In comparison with other participants, I understand the software content quickly (0.862)	Pintrich and De Groot (1990)	3.19	1.12
	SE2	In comparison with other participants, I am ahead in regard to the course content (0.848)	Pintrich and De Groot (1990)	2.81	1.12
	SE3	I know I have the necessary abilities to learn the course content (0.706)	Pintrich and De Groot (1990)	4.30	0.84
Technology Readiness $\rho_c = 0.870$ AVE = 0.696	TR1	I feel confident when using IT systems (0.908)	van der Rhee *et al.* (2007)	3.85	1.15
	TR2	I understand new high-tech products and can explain them (0.910)	van der Rhee *et al.* (2007)	3.63	1.21
	TR3*	New technology is often too complex to be useful (reverse coded)	van der Rhee *et al.* (2007)	2.06	1.08
	TR4*	Technology helps people to gain more control over their daily lives	van der Rhee *et al.* (2007)	3.52	1.07
	TR5	I like to be one of the firsts to be using new technologies (0.659)	van der Rhee *et al.* (2007)	3.20	1.25
	TR6*	Technology makes me more efficient	van der Rhee *et al.* (2007)	3.89	0.89

Subdimension	Item	Statement (loading)	Source	Mean	SD
Self-Regulated Learning $\rho_c = 0.840$ AVE = 0.637	SRL1	I check my learning success to make sure that I understand everything (0.804)	Pintrich and De Groot (1990)	3.71	0.96
	SRL2	I willingly study more than I am required to (0.829)	Pintrich and De Groot (1990)	2.78	1.24
	SRL3*	Before I start studying, I think about what I need to study	Pintrich and De Groot (1990)	4.08	0.97
	SRL4	I work hard to pass the course even if I do not like the content (0.760)	Pintrich and De Groot (1990)	3.66	1.09
Perceived Importance $\rho_c = 0.869$ AVE = 0.691	PIT1	I can easily tell where the participation in this course will be beneficial to me in future activities (0.893)	Developed during the first step of our research process	3.93	0.98
	PIT2	I am aware of the significance of the course in terms of how it will help me solve upcoming tasks (0.886)	Developed during the first step of our research process	3.97	0.96
	PIT3	The course contents are very significant for carrying out many tasks in a company (0.700)	Developed during the first step of our research process	4.03	0.91
	PIT4*	I have the feeling that it is important to my future employer that I complete this course	Developed during the first step of our research process	4.18	0.85
	PIT5*	I have the feeling it would have been a disadvantage for me to not have completed this course	Developed during the first step of our research process	3.70	1.15
Intrinsic Value $\rho_c = 0.807$ AVE = 0.515	IV1	I expect to learn exciting things in this course (0.827)	Pintrich and De Groot (1990)	3.86	1.00
	IV2	I expect to learn useful things in this course (0.756)	Pintrich and De Groot (1990)	4.48	0.78
	IV3	I expect to learn new things (0.653)	Pintrich and De Groot (1990)	4.67	0.65
	IV4	I expect to meet interesting people (0.616)	Pintrich and De Groot (1990)	2.70	1.24
	IV5*	I am sure that I will be able to keep up with the content of this course	Pintrich and De Groot (1990)	3.53	1.06
	IV6*	I will score high in this course	Pintrich and De Groot (1990)	3.34	1.05
Perceived TML Process Quality (reflective first-order, formative second-order construct)					
Company Support	CS1	I feel the company supports my participation in the training	Developed during the first step of our research process	4.15	1.02
	CS2*	I am adequately substituted during my absence for the course (side job, other responsibilities)	Developed during the first step of our research process	3.38	1.72
	CS3*	I can concentrate on the learning content without interruption during training (without external interruptions)	Developed during the first step of our research process	3.70	1.20
Interactivity $\rho_c = 0.877$ AVE = 0.642	Int1	I contribute to the course (i.e., by asking questions) (0.807)	Siau *et al.* (2006)	3.46	1.18
	Int2	I participate in the course discussions (0.790)	Siau *et al.* (2006)	3.37	1.25
	Int3	The atmosphere in the course is characterized by a lively exchange (0.773)	Siau *et al.* (2006)	3.43	1.21
	Int4	The layout of the course encourages me to interact (0.833)	Siau *et al.* (2006)	3.79	0.95
Learning Group $\rho_c = 0.864$ AVE = 0.614	LG1	I can contribute to the group (0.767)	Developed during the first step of our research process	3.52	1.16
	LG2	Mutual support in the group is beneficial (0.765)	Developed during the first step of our research process	3.91	1.05
	LG3	I have the feeling that I can give effective feedback to the group (0.826)	Developed during the first step of our research process	3.61	1.04
	LG4	I have the feeling that the group is capable of working efficiently (0.775)	Developed during the first step of our research process	3.76	0.95

Subdimension	Item	Statement (loading)	Source	Mean	SD
Quality of Exercises $\rho_c = 0.887$ AVE = 0.613	QE1	The employed exercises were very helpful (0.776)	Developed during the first step of our research process	3.65	1.11
	QE2	The exercises are easy to understand (0.655)	Developed during the first step of our research process	3.66	0.91
	QE3	The employed exercises help me gain a deeper understanding of the material (0.868)	Developed during the first step of our research process	3.74	1.08
	QE4	I understand the purpose behind the exercises carried out so far (0.827)	Developed during the first step of our research process	3.80	1.02
	QE5	The exercises can ensure the active participation of participants in the course (0.773)	Developed during the first step of our research process	3.71	1.09
IT Process Support $\rho_c = 0.911$ AVE = 0.673	PS1	The provided IT systems increased the communication possibilities between the participants and the trainer (0.775)	Developed during the first step of our research process	3.93	0.96
	PS2	The provided IT systems increased the possibility to interact with other participants (0.767)	Developed during the first step of our research process	3.58	1.20
	PS3	The provided IT systems provide a good structure of the process of the whole course (0.876)	Developed during the first step of our research process	3.91	0.96
	PS4	The provided IT systems made it possible for me to structure my learning process (0.896)	Developed during the first step of our research process	3.80	1.04
	PS5	The provided IT systems increase the possibility to follow the learning process (0.778)	Developed during the first step of our research process	3.86	0.95
Transparency of the Training Process $\rho_c = 0.864$ AVE = 0.561	TP1	The instructor ensures that I am informed about the next process step at all times (0.746)	Developed during the first step of our research process	4.19	0.80
	TP2	The instructor provides sufficient orientation aid so that I can position myself in the exercise process (0.730)	Developed during the first step of our research process	4.26	0.80
	TP3	The connection between the learning content and learning goals of the course is clear to me (0.679)	Developed during the first step of our research process	4.11	0.81
	TP4	The learning goals of the course were communicated clearly by the instructor (0.788)	Developed during the first step of our research process	4.32	0.70
	TP5	The connection between various contents and the learning goals of the course was made clear by the instructor (0.795)	Developed during the first step of our research process	4.30	0.70
Fit $\rho_c = 0.894$ AVE = 0.680	Fit1	The learning content was elaborated using examples from my areas of interest (0.819)	Developed during the first step of our research process	3.59	0.95
	Fit2	The course contents were tailored to my needs (0.880)	Developed during the first step of our research process	3.21	1.02
	Fit3	The employed IT tools matched my needs (0.840)	Developed during the first step of our research process	3.44	0.95
	Fit4	The support in the learning process matched my expectations (0.755)	Developed during the first step of our research process	3.63	0.92
Satisfaction ($\rho_c = 0.808$, AVE = 0.678)					
–	Sat1	I am very satisfied with the course (0.857)	Arbaugh (2001)	3.82	0.85
	Sat2	I would recommend this course to others (0.789)	Arbaugh (2001)	3.91	1.03

Subdimension	Item	Statement (loading)	Source	Mean	SD
Perceived Learning Success ($\rho_c = 0.871$, AVE = 0.628)					
–	LS1	My technical knowledge has definitely increased (0.792)	Arbaugh (2001)	3.94	0.90
	LS2	I can explain important terminology and facts from the course (0.831)	Arbaugh (2001)	3.99	0.80
	LS3	I can illustrate complex facts from the course (0.765)	Arbaugh (2001)	3.56	0.95
	LS4	I can identify contradictions and similarities in the learning content of the course (0.779)	Arbaugh (2001)	3.60	0.94

*Item was dropped during data analysis.

Appendix 2: Evaluation of the formative measurement models

We relied on the guidelines by Cenfetelli and Bassellier (2009) to evaluate our formative measurement models that are depicted in Table 4 (see pp. 19–20).

According to the first guideline, we checked for multicollinearity by computing the VIF. The results indicate that multicollinearity is not a problem in our study since the highest VIF value (1.851) is below the limit of 5 (Hair *et al.*, 2014).

In their second guideline, Cenfetelli and Bassellier (2009) state that a large number of indicators will cause many nonsignificant weights. Since we observed only one nonsignificant weight for perceived TML process quality, two nonsignificant weights for perceived structural quality, and three for predisposition quality (at the level of 0.05, marked with "n.s." in Table 4), and their inclusion is based on theory, we decided not to drop any indicators. This decision is based on the argument that this is the first study of its kind, and it should be checked whether this lack of significance could be observed in different studies before questioning the relevance of these indicators (Cenfetelli and Bassellier, 2009).

The third guideline deals with the co-occurrence of positive and negative weights. Due to the fact that no indicator with a negative weight was found to be significant, there was no need to worry about this point in our study (Cenfetelli and Bassellier, 2009).

Guideline four suggests that researchers should check the indicator loadings when observing indicators that have a low indicator weight. As a reason, Cenfetelli and Bassellier (2009) point out that the indicator could only have a small formative impact on the construct (shown by a low weight), but it could still be an important part of the construct (shown by a high loading). If this is the case, the indicator is important and should be included in the measurement model (Cenfetelli and Bassellier, 2009). Chin (1998) stipulates that a loading of 0.5 is weak but still acceptable, and Stevens (1992) recommends a cutoff criterion of 0.4. Checking the results presented in Table 4, we can see that the loadings of two indicators with nonsignificant weights are below this cutoff criterion. Nevertheless, since this is the first study of this kind and the inclusion of the indicator is based on a solid theoretical basis and our research process, we follow Cenfetelli and Bassellier's (2009) advice to not drop the indicator. If similar results can be observed in future studies, the indicator should be dropped, and the suitability of the theoretical basis should be investigated.

In the fifth guideline, Cenfetelli and Bassellier (2009) recommend testing for nomological network effects and constructing portability. They suggest comparing the factor weights of the indicators across different studies. Due to the fact that, to the best of our knowledge, ours is the first study applying a holistic TML performance evaluation approach for vocational software trainings, a comparison of factor weights across different studies is not possible.

The sixth guideline cautions that the indicator weights can be slightly inflated when using the PLS technique (Cenfetelli and Bassellier, 2009). Since we rely on the new PLSc algorithm (Dijkstra and Henseler, 2015), our indicator weights should not be inflated.

In summary, the evaluation of our formative measurement models shows that the models fulfill the requirements posed by the guidelines by Cenfetelli and Bassellier (2009).