The Data Base for Advances in Information Systems

A Design Theory for Certification Presentations

Sebastian Lins
Karlsruhe Institute of Technology

Jan-Michael Becker
BI Norwegian Business School

Kalle Lyytinen
Case Western Reserve University

Ali Sunyaev
Karlsruhe Institute of Technology

Date of Acceptance: 4/3/2022

Authors' comment: Please note, this prost-print version contains some formatting issues in tables and figures. We hope that the journal will revert them throughout the final editing.

Please cite this article as follows:

Lins, S., Becker, J. M., Lyytinen, K., Sunyaev, A. (Forthcoming). A Design Theory for Certification Presentations . *The Data Base for Advances in Information Systems*, In Press.

A Design Theory for Certification Presentations

Sebastian Lins
Karlsruhe Institute of Technology

Jan-Michael Becker
BI Norwegian Business School

Kalle Lyytinen
Weatherhead School of Management, Case Western Reserve University

Ali Sunyaev
Karlsruhe Institute of Technology

Acknowledgments

Financial support for this work by the German Research Foundation (DFG) through the granted research project "*A Decompositional Analysis of IT Certifications in Electronic Markets and their Effect on Platform Customer and Provider Perceptions*" (SU-717/10-1) is gratefully acknowledged.

Abstract

Prior information system research remains inconsistent of the effects of system certifications. In their current use, certifications are often reduced to graphical seals. This approach fails to incorporate detailed assurance information emanating from the certification process. To address this gap, we adopt a design science approach and deploy a four-phase research design to clarify how to design impactful IS certification presentations. First, we identify sources of users' limited understanding of seals and formulate a design proposal for a certification presentation by drawing upon the elaboration likelihood model. In the second phase, we formulate and validate a set of design meta-requirements and guidelines to improve certification presentation, using cognitive load theory and Toulmin's model of argumentation as kernel theories. In the third phase, new certification presentations that comply with the proposed guidelines are developed and evaluated for their effectiveness. We show that presentations that augment seal-based certification presentations with richer assurance information improve certification effectiveness. This increases users' assurance and trust perceptions when the presentations align with the users' cognitive information processing needs in ways that reduce their cognitive load and enhance argument quality of assurance information.

Keywords: Certification; Assurance Seal; Elaboration Likelihood Model; Design Science; Cognitive Load Theory; Toulmin's Model of Argumentation.

Introduction

Information system (IS) owners have developed diverse strategies over the past years to reduce users' uncertainty and anxiety surrounding unknown or ad hoc system use. Among these strategies, certifications by third parties are the longest and most pervasively used (Lins et al., 2020; Özpolat et al., 2013). Certifications are neutral third-party attestations of specific system characteristics, operations, and management principles to prove compliance with regulatory or industry requirements (Lansing et al., 2018). Well-known certifications include now among others *"Certified Privacy"* for webshops*, "CSA STAR"* for cloud services, and management standards for security "*ISO/IEC 27001—Information security management systems"*. System owners can benefit from presenting well-established and recognized certifications in their websites or system user interfaces because certifications increase users' *perceived assurance* concerning system use and foster users' *trust perceptions* toward the system owner (Löbbers et al., 2020). In the context of IS use, system owners typically signal the possession of such certifications by placing assurance seals as certification presentations on their websites or in their system interfaces (Figure 1). However, system owners witness that intended effects of certifications and corresponding seals often remain unfulfilled (Adam et al., 2020; Lansing et al., 2018; Löbbers et al., 2020), ultimately wasting financial and human resource for acquiring certifications (i.e., annual costs are usually between $2,500 and $75,000) and ongoing efforts to stay in compliance with IS certification requirements (Mavlanova et al., 2016).

Insert Figure 1 About Here

To account for this conundrum, past research has analyzed to what extent and how certifications affect users' perceptions and why these differential effects emerge (Lins & Sunyaev, 2017). Prior research concludes that certifications' ineffectiveness stem in part from users' limited understanding of the nature and role of certifications (Lowry et al., 2012). Users are typically poorly informed about the meaning of certifications and certification requirements (Kimery & McCord, 2006; McKnight et al., 2004; Portz et al., 2000; Statista Survey, 2017). Prior research posits that users' limited understanding likely originates from ineffective certification presentations and the extent to which assurance information is made accessible to users (e.g., Lansing et al., 2018; Lowry et al., 2012). As a consequence, several researchers have called for additional studies on how certification presentations can be designed to improve user's processing and understanding of certification information, thereby generating expected assurance and trust effects (e.g., D. J. Kim, Steinfield, & Lai, 2008; Kimery & McCord, 2006; Lansing et al., 2018; Lowry et al., 2012). To answer these calls and resolve the issues surrounding certifications' ineffectiveness, we focus on IS certification presentations and clarify how to improve their design. Thus, this study's central design artifact is an IS certification presentation scheme that provides the means to communicate certification-relevant information. We next examine the antecedents and effects of certification presentations on users' information processing and based on such analysis develop a design theory for IS certification presentations. We ask:

How can certification presentations be designed to increase users' perceived assurance and trust perceptions?

We follow a theory-driven design science approach (Arazy et al., 2010) to clarify how to design effective IS certification presentations. In line with design science canons (e.g., Kuechler & Vaishnavi, 2008), the study design is divided into four phases: (1) becoming aware of users' problem to understand certifications, (2) suggesting a theory-based design of presentations, (3) developing and evaluating presentation designs, and (4) drawing conclusions for the design of effective IS certification presentations. Each phase tackles a critical sub-question of the overall design study.

The remainder of the paper proceeds as follows. Next, we present an overview of certifications and related research. In the subsequent section, we briefly introduce our research context and design science research and outline our four-phased research design that aligns with the theory-driven design approach by Arazy et al. (2010). Results of the problem awareness, design, and development and evaluation phases are then reported covering research methods and findings. We conclude by discussing this study's principal findings, the implications of our findings for certification research and practice and by suggesting directions for future research.

Certifications in Information System Use

The Importance and Value of Certifications

IS certifications involve three actors: system owners, users, and certification authorities (Lins et al., 2020; Renner et al., 2021). System owners offer IS to (potential) users, are responsible for system operations, and acquire IS certifications. Users are individuals or organizations that rely on IS certifications to reduce their uncertainties concerning the qualities and features of the system. Certification authorities are independent, neutral intermediaries between users and system owners that provide forms of oversight to deter or punish inappropriate behavior by the system owner. The oversight covers, among other things: assessing the system documentation about its security and data protection measures, interviewing system owners' employees, or conducting on-site assessments or penetration tests to evaluate the system's compliance (Lansing et al., 2018). If the target system for a certification adheres to specified requirements, the certification authority awards a formal testimonial. The system owner is then permitted to communicate their certification compliance and corresponding certification information to users and outside stakeholders (e.g., business partners). Such certification information typically covers information about the certification target (e.g., a cloud service), certification's objectives (e.g., ensuring GDPR compliance), organizations that are involved in a certification's issuance and audit processes, and corresponding assurance information relevant for users' subsequent system use (Lansing et al., 2018).

The number and variety of such certifications have increased continuously as the IS use has diversified and expanded. Recently, the EU GDPR and EU AI Act have foreseen certifications as the primary mechanism for organizations to demonstrate compliance with GDPR and AI requirements across industries and legislative regions. Likewise, the EU Cybersecurity Act introduced an EU-wide cybersecurity certification framework to harmonize existing security certifications and foster the development of novel security certifications in emerging domains, such as distributed ledger technology or artificial intelligence. Certifications are valued, in particular, by small- and medium-size system owners in electronic markets (D. J. Kim et al., 2017; Sunyaev & Schneider, 2013). They often lack a strong, confidence-building market position and a strong reputation and seek to embed certification information in their system use as testimonies of proper legal, organizational, and technological protections for users. Organizations most commonly adopt certifications because they can increase users' *perceived assurance* and *trust perceptions*.

Perceived assurance refers to a user's perception of the likelihood that the system owner has in its interest to protect the user's confidential information and that it applies proper security measures such as authentication, encryption, and non-repudiation to achieve such protection effects (D. J. Kim, Ferrin, & Rao, 2008). The extent to which such protections are salient to users depends on how clearly they understand the levels of security and data protection measures. When users find system-related security and privacy information (e.g., a privacy policy, a security disclaimer, a safe usage guarantee, etc.) or information on protection mechanisms (e.g., encryption, authentication, SSL technology, etc.), they will likely recognize the owner's intent to protect the users' information. Users can also infer the better quality of the system and forecast its security and privacy as well as the integrity of related management operations with greater accuracy and confidence (Kimery & McCord, 2008).

Trust perceptions are defined as users' perceptions of the system owner's competence (ability to do what the user needs), benevolence (care and motivation to act in the user's interests), and integrity (honesty and promise-keeping) (McKnight et al., 1998). Trust often emerges as a critical element when users must deal with uncertain and uncontrollable situations. Certifications foster users' trust perceptions because certifications reflect and signal institution-based trust (McKnight et al., 2004). Institutional trust theory (Zucker, 1986) posits that prevailing institutional structures (e.g., guarantees, certifications, or third-party intermediaries) create an environment in which users can feel safe. Thus, such institutionally mediated structures foster the growth of trust and cooperation between the system owner and the user. Certifications can also provide strong trust-assuring arguments for system owners because they associate the system owner's claims of system features with an independent and trusted authority (K. Kim & Kim, 2011). To the extent that users perceive the authority as credible and trustworthy, the trust applied to this authority is transferred to the system owner, which is referred to as *trust transference* (Stewart, 2003).

Related Research on Certifications

The literature on IS certifications and related web seals constantly grew in recent decades, predominantly in three research streams.[1] First, various scholars have examined how to develop (e.g., Lynn et al., 2016) or innovate certifications and underlying attestation processes (e.g., Lins et al., 2019). Second, research taking an organizational perspective analyzes the motivations of system owners to adopt certifications, how they internalize certifications, and whether they can harness the benefits of certifications (e.g., Hsu et al., 2012). Third, research adopting a user perspective seeks to explain how certifications affect users, why these effects occur, and how to predict these effects (e.g., Löbbers & Benlian, 2019). We position our study in this user-centered stream of certification research because we adopt a user perspective to examine users' certification understanding problem and how to design IS certification presentations increasing their effects on users.

Extant research's evidence of the positive effects of certifications in increasing users' perceived assurance and trust perceptions remains inconclusive (Table 1; Adam et al., 2020; Löbbers et al., 2020). One group of studies has observed that certifications have a significant positive effect on users' *perceived assurance* (e.g., K. Kim & Kim, 2011; Yang et al., 2006) and *trust perceptions* (e.g., Hoffmann et al., 2015; Mavlanova et al., 2016). Another group found no significant positive effect on *perceived assurance* (e.g., Lowry et al., 2012; Rifon et al., 2005) or *trust perceptions* (e.g., D. J. Kim, Ferrin, & Rao, 2008; McKnight et al., 2004). Such inconsistencies cast doubt on the potential effectiveness of using certifications. To overcome this conundrum, research has started to examine contingency factors that shape certification's effects, including the effects of context (e.g., the structural elements of a certification; Lansing et al., 2018, and characteristics of websites on which the certification is embedded; Adam et al., 2020) and the role of user characteristics (e.g., national culture; D. J. Kim et al., 2017, or personality traits; Löbbers & Benlian, 2019).

Insert Table 1 About Here

More importantly, research also suggests that users often face difficulties in processing and understanding certification presentations (Kimery & McCord, 2006; Lowry et al., 2012; Lynn et al., 2016). Users are unfamiliar with certifications and tend to have a limited understanding of what the underlying assurance mean, or they simply misunderstand them. Given challenges in user's processing of certifications, several researchers have called for additional studies on boundary conditions that explain the effective use of certification presentations (e.g., D. J. Kim, Steinfield, & Lai, 2008; Kimery & McCord, 2006; Lansing et al., 2018; Löbbers et al., 2020; Lowry et al., 2012; Lynn et al., 2016). However, we currently lack knowledge about how to design certification presentations in ways that overcome users' limited understanding, can increase users' perceived assurance and convey trust (Lynn et al., 2016). Rare exceptions are our related research that examined how to design certification presentations by embedding peripheral cues (i.e., authority, social proof and liking cues) to achieve certifications' effects if users are unmotivated or incapable to read underlying assurances (Lins & Sunyaev, 2022). We thus require more knowledge on how to design certification presentations in ways which improves user's processing and understanding of assurances generating expected assurance and trust effects.

System owners, certification authorities, policymakers, and users will be detrimentally affected if IS certifications fail to achieve vital beneficial effects. Not only system owners' significant investments in IS certifications (i.e., usually between $2,500 and $75,000 per year) and their recurring efforts to maintain compliance might be wasted or ineffective in providing assurances to users (Mavlanova et al., 2016). But also, certification authorities' extensive efforts to develop novel IS certifications to answer regulatory changes (e.g., concerning the GDPR or EU AI Act) may not be worthwhile, threatening their business longevity in the long run. In a similar vein, ineffective IS certifications will demand policymakers to decree novel mechanisms to prove regulatory compliance, such as in the case of proving compliance with data protection requirements imposed by the GDPR. Users will be less able to reduce prevalent information asymmetry faced in electronic markets, hindering them from differentiating between high- and low-quality IS and making informed IS selection decisions, ultimately facilitating the emergence of a classical *"market of lemons"*(Akerlof, 1970). Consequently, it is crucial for IS researchers and practitioners to understand how users process and understand the nature and role of certifications and underlying assurances, and develop means to enhance the effectiveness of IS certification presentations leading to expected certification effects. We start to tackle this open challenge in this study. We next outline out study design, including the research context and our four-phase approach to clarify how to design certifications presentations to increase users' perceived assurance and trust perceptions.

Study Design

Research Context

Certifications are used across various application domains, such as the certification of a system owner's compliance with quality (i.e., *ISO 9001*), environmental (i.e., *ISO 14001*) and security management standards (i.e., *ISO/IEC 27001*), the *Payment Card Industry Data Security Standard* (PCI-DSS), or the *Health Insurance Portability and Accountability Act* (Danylak et al., 2022; Renner et al., 2021). In the IS discipline, the predominance of certification research examines certifications in the context of electronic markets that manifest great user uncertainty (Table 1; Löbbers et al., 2020). To position this study in related IS research on certifications, we selected electronic markets as a research context. In particular, we focus on cloud service markets and related certifications. We chose this specific context because cloud services are now widely used and have become a critical element of the encompassing IT infrastructure and related services (Benlian et al., 2018; Schneider & Sunyaev, 2016). Cloud service use permeates now almost every aspect of the everyday use of IT, including ecommerce, storage services (e.g., *Dropbox*), security (e.g., *Carbonite*), social media, etc. Cloud services have become highly popular in that they offer a large variety of daily IT services (e.g., online storage, office software, and collaboration tools), and such services must withstand an unexpected volume of online use. We deem the cloud service market as an appropriate context for our study because the related systems are characterized by a high degree of user uncertainty concerning system use and its operations (Lins et al., 2018; Trenz et al., 2018). Users cannot examine cloud services and their features in advance due to the lack of regulatory controls and transparency in relation to their security and privacy protection. Consequently, users find it difficult to determine which cloud system owners can be trusted to provide satisfactory service and which ones can protect users' personal and financial information. We chose cloud storage services (e.g., *Dropbox* or *Google Drive*) as an empirical setting because they all share common cloud service characteristics. For example, the privacy and security measures of such storage services cannot be evaluated in advance, and the services handle often significant amounts of sensitive user data (e.g., information obtained through backups and sharing). To manage the aforementioned challenges, a variety of certifications have recently been proposed to cloud markets, such as *EuroCloud's "StarAudit"*, *Cloud Security Alliance's "Security, Trust and Assurance Registry"*, and *Competence Center Trusted Cloud's "GDPR Cloud Certificate"* (Lins et al., 2016). While users may rely on these certifications when selecting a cloud service, they will have a limited understanding of the related assurances due to their novelty. Moreover, the variety and ambiguity of related certifications are likely to threaten their effectiveness (Lansing et al., 2019).

Four-phase Design Science Research Approach

We address concerns of certification presentation effectiveness by engaging in theory-driven design science research (DSR). In its essence, DSR involves the creation of new knowledge through design of novel artifacts and analysis of the use or performance of such artifacts to address unsolved problems or address them more effectively/efficiently than previous attempts, along with reflection and abstraction to improve and understand the behavior of IS's aspects (Hevner et al., 2004; Vaishnavi et al., 2017). DSR is considered as the link between research and practice (Peffers et al., 2007) and has been applied in diverse contexts, such as to guide the development of an evaluation framework for patterns (Petter et al., 2010), to design and evaluate systematic search systems (Sturm & Sunyaev, 2019), or to develop innovative certification processes (Lins et al., 2019).

Given the increasing interest in DSR in the IS discipline, there has been a continuous (and controversial) scientific discourse on what DSR is (e.g., Hevner et al., 2004) and is not (e.g., Baskerville, 2008), how to conduct DSR studies (e.g., Nunamaker et al., 1990; Peffers et al., 2007; Vaishnavi & Kuechler, 2015; Walls et al., 1992), recommendations and criteria for rigor, utility, and aesthetic (e.g., Baskerville et al., 2018; Venable et al., 2016), and guidelines on positioning and presenting DSR studies (e.g., Gregor & Hevner, 2013; Peffers et al., 2018), among others. On the contrary, this excess of advice and expectations for carrying out DSR also challenges current researchers, making it difficult to carry out DSR projects and leading to less research that actually applies DSR (Peffers et al., 2018).

We build on Peffers et al. (2018)'s definition of five genres in IS DSR to position our research in the comprehensive DSR domain. With this study, we strived to apply design research involving designs driven by supporting kernel theories from related disciplines, and therefore align with the genre *'IS design theory'* (Peffers et al., 2018). The IS design theory genre builds on kernel theories from the natural and behavioral sciences to justify the proposition of design-related hypotheses (Peffers et al., 2018). Developing IS design theories is at the core of this genre to generate and represent knowledge. Notably, prior work of Simon (1996), Walls et al. (1992), and Gregor and Jones (2007) inform this genre (Peffers et al., 2018). We decided to opt for a theory-driven design approach because grounding artifacts' design in kernel theories not only increases the designer's understanding of the problem domain but also helps formulate high-level design principles independent of technological constraints and specific implementation details (Arazy et al., 2010). Moreover, designs grounded in theory are expected to produce more effective IS since the laws of the natural and social world govern the components that comprise the system. By assigning this study to the IS design theory genre, we want to dissociate us particularly from the genres of *'Design science research methodology'* (e.g., Nunamaker et al., 1990; Peffers et al., 2007) and *'Design-oriented IS research'* (e.g., Österle et al., 2011; Winter, 2008), which emphasize the development of artifacts (e.g., systems, applications, methods) and not demand that a design is based on formal theory (Peffers et al., 2018).

In the light of the *IS design theory* genre, we apply the theory-driven DSR approach of Arazy et al. (2010) that builds on and extends the DSR approach by Walls et al. (1992), and we considered guidelines proposed by Gregor and Jones (2007) to develop design theories. Certification representations form the design artifact and are defined as a scheme that provides the means to communicate certification-relevant information. Existing examples of IS certification presentation include assurance seals, printed certificates, or specific websites comprising certification-relevant information. According to Arazy et al. (2010), five different components are required for constructing the new artifact: kernel theories, an applied theoretical model, meta-requirements, design guidelines, and testable design propositions. To ease readers' understanding and conform to prevalent DSR canons (e.g., Kuechler & Vaishnavi, 2008; Vaishnavi & Kuechler, 2015), the study design is divided into four phases: (1) improving awareness of users' problem of understanding certifications, (2) suggesting a theory-based design of presentations, (3) developing and evaluating the presentation design, and (4) drawing conclusions. Each phase tackles a critical sub-question of the overall design science study and comprises corresponding components of Arazy et al. (2010)'s theory-driven design framework (Table 2).

Insert Table 2 About Here

In the **problem awareness phase** (Vaishnavi & Kuechler, 2015), we examine the problem domain, identify the design problem (Arazy et al., 2010), and propose a revised certification presentation to address the prevalent issues in certification presentations. In particular, by analyzing how users process certification presentations we try to identify the origins of users' limited understanding of certifications. Here, we draw on a well-established dual-process theory (i.e., the ELM) that accounts for the presence and role of alternative forms of information processing in shaping human's beliefs, attitudes, and behavior (Evans, 1984; Sloman, 1996). In this study, using ELM as kernel theory helps framing and testing the proposed solution for certification presentation that improves user understanding.

In the **suggestion phase** (Vaishnavi & Kuechler, 2015), we formulate the requirements and guidelines for designing a certification presentation that increase users' perceived assurance and trust perceptions. Per the assumptions of the ELM, we consider two major determinants of how users will process certification presentations and related information. We note that a low *cognitive load* is required to process certification presentations and that high *argument quality* is consequently necessary for persuading users. We develop and test an applied theoretical model (Arazy et al., 2010), guided by the principles of cognitive load theory (Sweller et al., 1998) and Toulmin's model of argumentation (Toulmin, 1958) as kernel theories to improve users information processing and design certification presentations that promote such goals. Building on our theoretical model and related literature, we derive meta-requirements and design guidelines (Arazy et al., 2010) that serve as input to formulate a design theory guiding designers in creating more effective certification presentations.

In the **development and evaluation phase** (Vaishnavi & Kuechler, 2015), we develop four instantiations of certification presentations and derive testable design propositions (Arazy et al., 2010) to evaluate the validity of the proposed design guidelines for improving the design. We implement proposed certification presentations in a controlled use setting and subject the users to an experiment that evaluates the effectiveness of implemented certification presentations in reducing the cognitive load and increasing argument quality related to certification presentation and to what extent they improve user's assurance information and trust.

Finally, in the **conclusion phase** (Vaishnavi & Kuechler, 2015), we consolidate the design results in a design theory for certification presentations, discuss contributions for research and practice, and point out limitations of our design approach to guide future research.

Table 3 summarizes relevant terms and their definitions. Figure 2 summarizes the objectives of each phase, used kernel theories, the applied research methods and their respective outputs, and how the findings of each phase informed the following phase. In the following sections, we will provide a summary of each phase.

Insert Table 3 About Here

Insert Figure 2 About Here

Problem Awareness Phase: Why Do Certifications Fail in Achieving Effects

Understanding the Problem Domain

First, we examined the problem domain and reviewed extant research on IS certifications to identify the origins of ineffective certification presentations and the corresponding design problem (Arazy et al., 2010). The output of this phase is a design proposal for the following research efforts (Vaishnavi & Kuechler, 2015). Delving into the problem domain reveals that not only the importance and number of IS certifications steadily increases but so do the issues and concerns of users (Table 4). Users share incorrect or uncertain answers when being asked about the meaning of the displayed certification. They are also unaware of what a system owner had to do to acquire the certification (e.g., Löbbers et al., 2020; McKnight et al., 2004; Moores, 2005; Portz et al., 2000). This situation hampers the effectiveness of the certification and decreases users' perceived assurance and trust perceptions. Furthermore, research reveals that users' recollections of offered certifications are typically shallow. Users are not familiar with the related certification requirements (Kimery & McCord, 2006; Löbbers et al., 2020) and experience expectation gaps between the assurances the certification intends to provide and the system qualities users perceive (e.g., Lansing et al., 2019; Odom et al., 2002).

Reflecting these issues surrounding IS certifications reveals that IS certification presentations fail in effectively communicating certification-relevant information. This is in line with prior research propositions, claiming that users' limited understanding likely originates from ineffective certification presentations and the extent to which assurance information is made truly available (e.g., Lansing et al., 2018; Löbbers et al., 2020; Lowry et al., 2012).

Insert Table 4 About Here

Dual-process theories offer promising clues about why certification presentations are likely to fail in communicating their intended content. Dual-process theories suggest that processing external information is the primary driver in the user's attitude change and related behavioral change (Eagly & Chaiken, 1993). In this study, we draw upon one of the most well-known dual-process theories called the ELM (Petty & Cacioppo, 1986) as our kernel theory (Arazy et al., 2010). Past research has found substantial support for the theory's predictive and explanatory power in a variety of behavioral domains (e.g., Angst & Agarwal, 2009; Bhattacherjee & Sanford, 2006; Mak & Lyytinen, 1997). More importantly, the ELM provides a theoretical foundation for differentiating how users process graphical cues (i.e., assurance seals as certification presentations) versus arguments contained in background texts (e.g., certification-relevant assurance information). Thus, the model offers a way of framing a solution for presenting certifications more effectively to users.

According to the ELM, information recipients (i.e., users) vary in their abilities and motivations to elaborate on external information (i.e., certifications), which constrains how subsequent information processing will impact users' attitude formation (Petty & Cacioppo, 1986). The ELM argues individual judgments are not always based on the effortful processing of external information. Instead, they are sometimes based on the less-effortful processing of peripheral cues. The ability and motivation for processing are captured in the *"elaboration likelihood"* construct. The term *“elaboration”* suggests people add something of their own to information provided in the communication, taking it beyond a mere verbatim encoding. People in a high elaboration likelihood state are likely to scrutinize the information message (referred to as cognitive processing). Therefore, they tend to be more persuaded by argument quality than by peripheral cues. In contrast, those in a low elaboration likelihood state lack the motivation or the ability to deliberate thoughtfully and are motivated to follow simple heuristics and peripheral cues in processing information (referred to as peripheral processing).

In IS use, certifications are typically presented in the form of assurance seals embedded on a website or system interface (Figure 1; Löbbers et al., 2020). The seals convey graphical cues regarding the underlying information and can be considered external information users will process to evaluate, among other things, the system owner’s trustworthiness. Previous research shows that seals act mainly as peripheral cues and trigger simple heuristics because of the presence of plain graphical presentation (D. J. Kim et al., 2017; Lowry et al., 2012; Yang et al., 2006). For example, if users recognize a certification seal in an online shopping system, a simple heuristic is triggered, and they perceive the subsequent shopping experience in the system as safe and secure (Yang et al., 2006). Consequently, seals affect users’ future attitudes and behaviors (i.e., *assurance perception* and *trust in a system owner)* by promoting peripheral processing.

People generally engage in peripheral processing in low-involvement situations in which they are neither capable of processing information thoroughly nor motivated to do so. They will engage in cognitive processing when they attempt to consider carefully the true merits of the presented information. These are called high-involvement situations (Petty & Cacioppo, 1986). Cognitive processing requires individuals to think critically about related arguments typically couched in a persuasive message (e.g., in the case of certifications, these are related security- and privacy-assuring statements). We posit that currently used certification presentations fail to trigger or enable such cognitive processing among users because the certification presentations have been reduced to graphical cues (i.e., seals) that do not incorporate comprehensible assurance information (e.g., with respect to underlying assurance requirements that a system owner must fulfill). This claim is backed by past research showing that users are often unfamiliar with certification seals and tend to either have a limited understanding of related assurances or to misunderstand them (Table 4; e.g., Kovar et al., 2000; Löbbers et al., 2020; McKnight et al., 2004; Moores, 2005; Portz et al., 2000). Consequently, users who are motivated and who can engage in cognitive processing (i.e., have a high elaboration likelihood) now lack detailed information to reason and change their attitude (and consequently their behavior) toward the system owner because they only have access to information that allows peripheral processing. We posit that these issues with certification presentation threaten—or at least diminish—certifications’ effectiveness.

Per the ELM, we argue that certification presentations during IS use need to support both cognitive and peripheral information processing, whereby users can choose the extent to which they elaborate on the provided information individually depending on the situation. Thus, some users may seek out arguments related to certification information, whereas others are likely to remain reliant on peripheral cues. The basis for this argument is that elaboration likelihood is not an automatic personality trait determining individual differences in processing such information. Rather, the state varies by context for the same individual (Bhattacherjee & Sanford, 2006; Lowry et al., 2012). Consequently, we propose certification presentations should comprise a graphical seal serving as a peripheral cue (facilitating peripheral processing) and detailed assurance information that richly informs users of underlying assurances (facilitating cognitive processing). Such certification presentations enable system owners to address cases in which users process external information through either high or low elaboration likelihood states. More importantly, these types of certification presentations then ensure that users’ attitudes will change resulting in behavioral change. We next elaborate on whether our design proposal of a certification presentation comprised of a seal and detailed assurance information will have an impact on certifications’ effectiveness by developing and testing hypotheses.

Hypotheses Development

Users can make better-informed decisions when they are presented with embedded assurance information in the certification presentations in ways that are highly structured, up to date, issued by an independent certification authority and inclusive of detailed information about certification requirements and related assurances. Accurate and complete assurance information embedded in a certification presentation will overcome the seals’ shortcomings (i.e., limited understanding and expectation gaps) and enable users to understand a larger amount of assurance information which prevents misinterpretation. Such information enables users to ascertain information that is more reliable about system owner’s behaviors. Related research has shown that educational interventions informing users about the assurances the seal provides influence user awareness and improve the perceived importance of seals (D. J. Kim, Steinfield, & Lai, 2008). In contrast to graphical seals, assurance information, such as information on applied security and data protection safeguards can be used (in a high elaboration likelihood state) by users to think critically about the system owner's efforts to protect system behaviors thereby increasing users' perceived assurance. In addition, users can rely on the information to verify whether the system owner complies with the certification requirements. Certifications can then reduce prevalent information asymmetry as the key issue preventing users to develop online trust (Lynn et al., 2016). We posit the following hypothesis:

> Hypothesis 1 (H_1): A certification presentation comprised of a seal and detailed assurance information increases users' perceived assurance when compared to certifications presented only by seals.

Users are not only uncertain when questioned about the meaning of displayed certification seals but are also unaware of what a system owner must do to acquire a certification (e.g., McKnight et al., 2004; Moores, 2005). Providing assurance information about the certification itself (i.e., the validity period or the certification scope), the certification process (i.e., system owner self-assessments or independent audits), and the issuing certification authority (i.e., whether the certification authority is accredited and reputable) can counteract this shortcoming (Lansing et al., 2018). More importantly, scrutinizing this information will impact users' perceptions of the issuing certification authority, thereby strengthening the trust transference. Assuming a certification authority is trustworthy, a certification can establish a cognitive association between a certified system owner and a certification authority whereby a user's trust in a certification authority is transferred to a system owner (Hu et al., 2010; K. Kim & Kim, 2011; Stewart, 2003). In contrast to graphical seals, assurance information, such as information about a certification authority's attestation practices, can be used by users (in a high elaboration likelihood state) to validate certification authorities' trustworthiness, thereby strengthening the trust transference from the certification authority to the certified system owner. We posit the following hypothesis:

> Hypothesis 2 (H_2): A certification presentation comprised of a seal and detailed assurance information increases users' trust perceptions when compared to certifications presented only by seals.

Experiment to Validate Certification Presentation Effects

We conducted a between-subject experiment using three groups of subjects to test the hypotheses. We asked subjects to evaluate their usage of a cloud storage system named "*MyCloudDrive*". See Appendix A for detailed information about the experiment, including details on the experiment design and corresponding screenshots, measurement, and data collection and analyses.

Experiment Treatments. We created three between-subject treatments and randomly assigned subjects to one of these treatments. First, a control group with *MyCloudDrive*'s website only**,** informing subjects about the cloud storage system's main features and characteristics. Second, a treatment group with *MyCloudDrive*'s website including a certification as a seal (i.e., *Symantec* seal) embedded into the website page. We expected it to trigger users' peripheral processing. Finally, a third treatment group with *MyCloudDrive's* website including a certification presentation comprising a seal and assurance information. In particular, we developed a detailed assurance text that provides subjects detailed information about the certification process, including information about the certification authority, the attestation process and the assurance statements, which all would facilitate users' cognitive processing.

Measurement. The experiment used validated scales from the literature and seven-point Likert-type scales to measure constructs. We measured users' perceived assurance (D. J. Kim, Ferrin, & Rao, 2008) and initial trust perceptions (McKnight et al., 2002), controlled for users' cloud knowledge (Flynn & Goldsmith, 1999), cloud involvement (Laurent & Kapferer, 1985), disposition to trust (Gefen, 2000), and privacy concerns (Buchanan et al., 2007).

Data Gathering. We recruited 300 participants from Amazon Mechanical Turk. We removed seven responses because the subjects failed attention checks or rushed through the experiment, resulting in 293 responses fulfilling initial data quality requirements. The subjects who participated in our experiment were nearly evenly distributed by gender (47% females); were, on average, 38 years old (minimum 19 years, maximum 63 years); and held a high school (12%), college (24%), undergraduate (51%) or postgraduate degree (10%). The subjects were mostly employed (81%), not looking for work (5%), or retired or unable to work (5%).

Data Analysis. First, we assessed the measurement model, demonstrating internal consistency, convergent validity, and discriminant validity. Second, both procedural and statistical methods (e.g., common latent factor (CLF) test and a zero constraints test using IBM SPSS Amos) were used to control for common method bias (CMB), applying the recommendations of Podsakoff et al. (2003). Third, we controlled for whether users' cloud knowledge, cloud involvement, disposition to trust, or privacy concerns differed across treatments to check for the presence of confounding effects and to determine whether the treatment groups were structurally equivalent. We did not find any significant differences for our control variables *cloud knowledge* and *involvement.* However, we did find significant differences in users' *privacy concerns*, which varied across the seal only and seal and assurance information treatment groups, with respondents having higher concerns in the seal and assurance information group (mean increment of .441; $p = .04$). In addition, we find a significant difference in the *disposition to trust* between the website and seal only treatment groups, with respondents having a higher disposition to trust in the seal only group (mean increment of .433; $p = .03$). We believe that these differences in our controls do not distort our findings. First, although concerns are higher in the seal and assurance information treatment, perceived assurance and trust perceptions are still higher compared to the seal and website treatment group. Therefore, one might assume that displaying assurance information is effective even in cases where users have high concerns. Second, although the disposition to trust is higher in the seal only treatment, the trust perception toward the system owner is not significantly different, which emphasizes the ineffectiveness of showing only the seal. Finally, we used multivariate analysis of variance (MANOVA) to test our hypotheses. The findings revealed a main effect of our manipulations on *perceived assurance* ($F(2, 290) = 14.45$, $p < .001$) and *trust perceptions* ($F(2, 290) = 4.03$, $p < .05$). Participants in the seal and information assurance condition had a higher *perceived assurance* and *trust perceptions* than those in the website condition ($F_{\text{assurance}}(1, 290) = 24.82$, $p < .001$; $F_{\text{trust}}(1, 290) = 7.41$, $p < .01$) and, more importantly, than those in the seal only condition ($F_{\text{assurance}}(1, 290) = 18.07$, $p < .001$; $F_{\text{trust}}(1, 290) = 4.25$, $p < .05$), **supporting hypotheses H_1 and H_2**. Participants in the website and seal only conditions did not differ significantly from each other in either of the two variables ($F_{\text{assurance}}(1, 290) = .57$, $p = .450$; $F_{\text{trust}}(1,290) = .46$, $p = .501$).

Discussion

We found support for our claim that a certification presentation comprising a seal and assurance information increases the *perceived assurance* and *trust perceptions* of users. Despite our careful selection of a seal and the prominent positioning of the seal on the website, we did not identify any significant effect from *only* displaying a seal. To analyze whether the participants did not notice the seal during our experiment, thus causing a non-significant effect, we complemented our experiment with another treatment condition with 100 subjects from Amazon Mechanical Turk. After showing the subjects the cloud service website, we asked whether the subjects noticed the seal and whether they could remember which seal was embedded. Subsequently, we called every participant's attention to our seal by showing them an additional reminder that *MyCloudDrive* had been awarded a particular certification seal. Our findings revealed that 65% of the subjects noticed the seal, and 68% of the subjects who noticed the seal could remember it correctly. Seventeen percent of the subjects did not remember the seal, and 12% of them remembered it incorrectly (i.e., they identified a different seal). These findings are in accordance with previous research on users' limited recognition and understanding of seals (e.g., D. J. Kim et al., 2017; Moores, 2005). More importantly, additional data analysis shows that, even with the forced seal cognition, the findings do not reveal significant effects of solely displaying a seal. The results strengthen our conjecture that, in addition to selecting the most impactful seal and forcing users to recognize this seal, users search for reliable and detailed assurance information, which achieves the intended effect of certifications. See Appendix A5 for results of our forced seal analysis.

Overall, our problem awareness phase highlights that our novel design proposal□a certification presentation comprising a seal and assurance information□may overcome the current shortcomings of certification presentations. Nevertheless, related research has taught us to be cautious about website design features' impact on trust perceptions and the effects of providing assurance texts. They may easily fail in achieving their intended effect, if assurance texts exhibit low argument quality (e.g., D. Kim & Benbasat, 2006; Lowry et al., 2012) or poor readability (e.g., Singh et al., 2011). Consequently, it is crucial to carefully design certification presentations that contain assurance information in a structured and comprehensible format as to prevent users' cognitive overload. While prior research has focused on analyzing just the design of seals (e.g., Aiken & Boush, 2006; D. J. Kim, Steinfield, & Lai, 2008; Moores, 2005), we next suggest how to design certification presentations that can convey proper assurance information in a format that can foster the certification's intended effects.

Suggestion Phase: Design of Effective Certification Presentations

Deriving Meta-Requirements

Following Arazy et al. (2010), we next derive meta-requirements to design certification presentations that convey assurance information. Meta-requirements specify a class of goals to which the design theory applies (Arazy et al., 2010; Walls et al., 1992). To derive meta-requirements, we formulate an applied theoretical model that "is an explanatory theory that borrows from the generic kernel theories and is formulated as a behavioral framework, yet it is informed and constrained by the design requirements" (Arazy et al., 2010, p. 461). This component was inserted by Arazy et al. (2010) to extend the original design science framework of Walls et al. (1992) since their framework provides very little direction on how the linkage between kernel theory and meta-requirements could be achieved. Figure 3 illustrates our proposed model. Per the ELM, two major determinants influence how users cognitively process assurance information: 1) the level of the cognitive load imposed by the assurance information (Petty & Cacioppo, 1986) and 2) argument quality that determines the subsequent level of user persuasion (Petty et al., 1981). In the following, we describe how these determinants impact certification presentations and how they can be manipulated in light of chosen theories. The Appendix B1 offer further information on kernel theories.

Insert Figure 3 About Here

The Impact of the Cognitive Load on Certifications' Intended Effects

When the processing demands of the elaboration task exceed the user's contextual processing capacity, cognitive overload results (Sweller et al., 1998). Cognitive load is referred to as the level of cognitive effort necessary to scrutinize and elaborate on the assurance information and related argument. In situations of cognitive overload, users' performance degrades and it hampers their understanding of information (Paas et al., 2004; Sweller, 1988). Subsequently, conditions that can reduce the cognitive load to manageable levels need to be established to ensure user's adequate performance and to improve her understanding of the information. Here, we build upon cognitive load theory (CLT) (Sweller, 1988; Sweller et al., 1998) because this theory focuses on the limitations of users' working memory during the performance of a complex cognitive task and has been applied frequently in IS research.

CLT's basic premise is that users' cognitive processing is heavily constrained by their limited working memory which can only process a limited number of elements at a time (Sweller et al., 2019). An element is anything that needs to be learned, such as assurance statements contained in certifications. Cognitive load is highly dependent on the level of element interactivity (Sweller, 2010). Information exhibiting low element interactivity allow individual elements to be learned independently of each other and so impose a low working memory load. In contrast, high element interactivity information consists of elements that heavily interact and so cannot be learned in isolation. Thus, the more elements that interact, the heavier the cognitive load. For example, when dealing with equations, all the elements associated with an equation must be considered simultaneously because all the elements interact, leading to high cognitive load. Users will then allocate working memory resources as much as possible to deal with those interacting elements. If cognitive load exceeds users' available working memory resources, it hampers learning (Sweller et al., 2019).

In the context of certifications, we assume that a wealth of assurance information and related arguments needs to be evaluated simultaneously to create a holistic sense of assurance. More importantly, assurance arguments are characterized by high element interactivity because arguments interconnect and consist of complex related statements (e.g., technical arguments describing implemented security mechanisms) (Lansing et al., 2019). Therefore, a high cognitive load will accrue from the inadequate structuring and presentation of assurance information. This high cognitive load will hamper the understanding of information and reduce users' perceived assurance. Subsequently, we posit the following hypothesis:

> Hypothesis 3a (H_{3a}): An increased level of the cognitive load has a negative impact on users' perceived assurance.

Providing rich assurance information about the certification and its authority will impact users' perceptions of the issuing certification authority, thereby strengthening trust transference. If users perceive a certification authority as trustworthy, then their trust in a certification authority is transferred to a system owner (K. Kim & Kim, 2011; Stewart, 2003). In the context of IS certifications, embedding assurance information can be used to express that a certification authority is an independent expert having high competence, which is an important antecedent of trustworthiness (McKnight et al., 1998). For example, assurance information should highlight that certification authorities are not primary subjects in relationships between system owners and users, but act as an intermediary to facilitate the system use by validating system owner's capabilities, possession of specific practices, or use of specific technologies that enable secure usage. However, providing rich (and interconnected) information on the certification and certification authority carries the risk of overloading users' cognitive resources. In particular, because casual users are unfamiliar and lack experience with certification authorities (McKnight et al., 2004; Moores, 2005), users cannot rely on long-term memory and related cognitive schemas to ease information understanding

(Sweller, 1994). If users inadequately process certification information due to a high cognitive load, users will be unable to elaborate on the certification's trustworthiness. In consequence, trust transference from the certification authority to the certified system owner is weakened. We posit:

Hypothesis 3b (H_{3b}): An increased level of the cognitive load has a negative impact on users' trust perceptions.

The Impact of Argument Quality on Certifications' Effects

The cognitive route of attitude change has been traditionally operationalized in ELM research by using as a measure argument quality (e.g., Angst & Agarwal, 2009; Bhattacherjee & Sanford, 2006). Argument quality refers to a subject's perception that a message's arguments are strong and cogent, as opposed to being weak and specious (Petty & Cacioppo, 1986). When subjects process information, argument quality becomes a strong determinant of persuasion because the cognitive route is executed. A strong argument generates favorable thoughts toward the target of arguments (Angst & Agarwal, 2009). However, conceptualizations of argument quality have been inconsistent in the prior literature. Some studies refer to argument quality as a measure of message valence (Mongeau & Stiff, 1993), some refer to argument strength (Petty & Wegener, 1999), while some refer to both (Areni & Lutz, 1988). When manipulating an argument's valence and strength, one must alter the argument's content by replacing phrases and words with stronger and more consistent ones. Consequently, providing design guidelines independent of the content of actual assurance information is difficult. Therefore, we next follow D. Kim and Benbasat (2006), who applied Toulmin (1958)'s model of argumentation to express the quality of an argument. Grounded on Toulmin's model of argumentation, we show that trust-building around certification presentations is achieved by structuring arguments around assurance information effectively as to convince the user of the quality of the information.

Empirical evidence suggests also that applying Toulmin's model for argument structuring affects users' beliefs. Proper and logical structuring has been shown to improve the level of the persuasiveness of respective arguments in the context of expert systems (Ye & Johnson, 1995) and online communities (Stegmann et al., 2007), among others. In line with these findings, we posit that increasing the argument quality of assurance information embedded in certifications will increase the users' perceived assurance. For example, if assurance statements about applied technical safeguards in the system exhibit a strong argument, they will generate thoughts that are more favorable and that persuade the user to feel more secure. We hypothesize as follows:

Hypothesis 4a (H_{4a}): Improved argument quality has a positive impact on users' perceived assurance.

In particular, D. Kim and Benbasat (2006) demonstrate that structuring the content of statements according to Toulmin's model will increase users' trust in the system. We argue alike that embedding assurance information related to the certification and its authority in ways that exhibit strong argument quality will increase users' trust perceptions of the certification authority. As such, providing detailed arguments and evidence for why a certification authority is perceived as competent increases users' trust perceptions. In general, performing certification attestations is a complex procedure requiring a certification authority to be organized and armed with a body of knowledgeable and experienced employees (Lins et al., 2020). Certification authorities with high competence are likely to provide better assessments of system owners. For example, qualified authorities would be likely to detect possible security gaps when checking systems. Providing arguments about a certification authority's competence, benevolence, and integrity can engender in users a willingness to depend on the certification authority (Lansing et al., 2018; McKnight et al., 1998). In consequence, higher argument quality strengthens the trust transference. We propose:

Hypothesis 4b (H_{4b}): Improved argument quality has a positive impact on users' trust perceptions.

Validating the Applied Theoretical Model

We conducted a survey among 352 participants to validate the applied theoretical model. See Appendix B for detailed information about the survey, including details on the survey design, measurement, and data gathering and analyses.

Survey Design. We built on the previous experiment and asked subjects to evaluate *MyCloudDrive* as an option for their personal use. We presented the subjects a website comprising assurance information related to the certification and asked the subjects to scrutinize the presented information.

Measurement. We reused the measures from our first experiment, measured argument quality (D. Kim & Benbasat, 2006), and build on NASA's Task Loading Index (Hart & Staveland, 1988) and the related formative rating scale to measure users' cognitive load.

Data Gathering. We recruited 352 participants from Amazon Mechanical Turk and omitted 24 responses due to survey misuses (i.e., failed attention checks, did not read the assurance information, or rushed through the survey), resulting in a sample of 328 responses that fulfilled the initial data quality requirements. Subjects participating in the survey were equally distributed by gender (female 48%); were, on average, 37 years old (minimum 19 years, maximum 72 years); and had a high school (18%), college (22%), undergraduate (50%) or postgraduate degree (8%). Further, the subjects were employed (75%), not looking for work (7%), looking for work (7%), or retired or unable to work (3%).

Data Analysis. The survey results were used to validate both the construct measurement scales and the proposed theoretical relationships. Analyses validated internal consistency, convergent validity, discriminant validity, and a full collinearity test provides evidence that CMB is not a major issue in our data. Since the cognitive load construct uses a formative construct specification, we checked for indicators' relative contributions and assessed calculated indicator weights and loadings as well as their significance. We dropped three cognitive load items because their relative and absolute contributions did not meet the required thresholds. Further, all variance inflation factor (VIF) values of the remaining cognitive load indicators were less than a threshold value of 3.3. Hence, the formative measurement model indicated low collinearity. To test our theoretical model and hypotheses, we used partial least squares structural equation modeling (PLS-SEM) and SmartPLS software, version 3.2.8 (Ringle et al., 2015). The significance of the structural path estimates was assessed using bootstrapping with 5,000 subsamples and with bias-corrected and accelerated confidence intervals (Ringle et al., 2015). We decided to rely on PLS-SEM instead of covariance-based SEM (CB-SEM) because we included cognitive load as formative measurement and PLS-SEM is the preferred approach to handle reflective and formative measurement models (Hair et al., 2019). While in general formative measures can also be used with CB-SEM, it often requires construct or model modifications to meet identification requirements (Hair et al., 2017). In our case, the theoretical model is not identified in CB-SEM and would require implausible model modifications to make it estimatable so that we use PLS-SEM which is able to estimate the model without modifications[2]. We tested the structural model by evaluating the direct effects and the explained variances (R^2). In addition, we controlled for several factors including *cloud service involvement*, *cloud service knowledge*, *privacy concerns*, and *disposition* to trust. Figure 4 summarizes the direct effects of the cognitive load and argument quality on perceived assurance and trust perceptions.

Insert Figure 4 About Here

The analysis results show the *cognitive load* negatively influences users' *perceived assurance* (direct effect: -.140; p-value < .01; confidence interval [-.236; -.051]) and *trust perceptions* (-.183; p < .001; [-.287; -.092]), **supporting H_{3a} and H_{3b}**. We thus can derive the following meta-requirement: *reduce the cognitive load imposed by the assurance information embedded in certification presentations to enhance the level of perceived assurance and trust perceptions (MR1)*. Similarly, the degree of *argument quality* positively influences users' *perceived assurance* (.460; p < .001; [.344; .559]) and *trust perceptions* (.538; p < .001; [.433; .630]), **supporting H_{4a} and H_{4b}**. Consequently, we conclude that a certification presentation should meet the following meta-requirement: *enhance the argument quality of the assurance information embedded in certification presentations to enhance the level of perceived assurance and trust perceptions (MR2).* Table 5 summarizes meta-requirements. Examining the effect-sizes f^2 (Hair et al., 2017) reveals that *argument quality* has a large effect on *trust perceptions* (f^2 = .402) and a medium effect on *perceived assurance* (f^2 = .261). In contrast, cognitive load has only a small effect on both *trust perceptions* (f^2 = .053) and *perceived assurance* (f^2 = .027). Effect-sizes therefore provide first indication that argument quality and thus MR2 might be more relevant for certification presentation designs than cognitive load and MR1 respectively.

Insert Table 5 About Here

Deriving Design Guidelines

Next, we derive design guidelines that characterize a family of artifacts meeting the corresponding meta-requirements (Arazy et al., 2010; Walls et al., 1992). Design guidelines address the design of a class of artifacts (e.g., different types of certifications), rather than a single artifact (e.g., a single certification by an individual certification authority). In the following, we outline design guidelines helping to meet the meta-requirements for certification representation design derived from CLT and Toulmin's model (Table 6).

Insert Table 6 About Here

Design Guidelines to Reduce the Cognitive Load

CLT distinguishes between three types of cognitive loads: intrinsic, extraneous, and germane (Sweller et al., 1998). The load is called *"intrinsic"* if it is imposed by the natural complexity of information that must be understood (Sweller et al., 1998). Information is more complex and thus more difficult to understand when it consists of multiple interacting elements that cannot be learned in isolation and are readily held in the working memory. Hence, a high number of information elements and their interactivity impose a high intrinsic cognitive load (Sweller, 2010). In the context of certifications, a high intrinsic cognitive load is likely to occur because assurance statements involve high element interactivity, and many, if not all, assurance elements must be processed simultaneously.

Intrinsic cognitive load only can be altered by changing the nature of what is learned, by the act of learning itself, or by changing the expertise of the subject (Sweller, 2010; Sweller et al., 2019). Researchers have frequently recommended managing the intrinsic load by reducing element interactivity between the information elements through segmentation procedures (e.g., Mayer & Moreno, 2003; Pollock et al., 2002), later becoming termed the isolated-interacting elements effect (Sweller, 2010; Sweller et al., 2019). Although full simplification of the certification information may not be possible, designers can allow users to digest one chunk of the presented information at a time before moving to the next element through segmentation (Mayer & Moreno, 2003). Designers should artificially isolate information elements by showing only a limited number of elements at a time, by providing hyperlinks to others, and by allowing subjects to choose freely which information to reveal next, if necessary (Pollock et al., 2002; Sweller, 2010). Users can then learn first the individual elements, followed by the interactions between these elements (Sweller, 2010). By delaying learning how elements interact, working memory is freed, facilitating learning compared to users required to both learn individual elements and how they interact, simultaneously. Building on this isolated-interacting elements effect, we propose that a certification presentation needs to fulfill the following design guideline: *manage the intrinsic cognitive load by reducing element interactivity using segmentation (DG1).*

If the load is imposed by the manner in which the information is presented to users, it is called *"extraneous"* (Sweller et al., 1998). Especially layout, structure, orientation, and navigation impose an extraneous cognitive load when users read and elaborate content online (DeStefano & LeFevre, 2007). The more resources users devote to dealing with extraneous load the less will be available for dealing with intrinsic load and so less will be learned (Sweller et al., 2019). Consequently, we propose that appropriate certification presentation designs should help decrease the extraneous cognitive load to free users' working memory (Sweller et al., 1998).

CLT is primarily concerned with techniques designed to reduce extraneous cognitive load, and hence, multiple approaches to reduce such an extraneous load have been proposed over time (Mayer & Moreno, 2003; Sweller et al., 2019; Van Merriënboer & Sweller, 2010). First, the split attention effect must be mediated (Sweller, 1988). Such an effect occurs when users are confronted with multiple sources of information, distributed either in space (spatial split attention) or in time (temporal split attention), which must be simultaneously processed. For example, a short assurance claim and a hyperlink to detailed support for the assurance claim can be displayed to a user. To understand the assurance claim, the user must now follow a hyperlink and will be redirected to another page. These multiple sources of information need to be replaced with one integrated source of information (Sweller et al., 1998). To counteract the split attention effect and to reduce the extraneous cognitive load, the number of hyperlinks must be reduced by using semantically meaningful hierarchical navigation structures (DeStefano & LeFevre, 2007). In addition, techniques must be used that do not change the main page or redirect subjects (i.e., by using textual overlays, hovering windows, frames, or rollover windows) when displaying ancillary and necessary material, such as definitions, commentary, and examples (Britt & Gabrys, 2002). Consequently, we argue that designers need to *mediate the split attention effect using a hierarchical structure, reduce the number of hyperlinks and prevent the redirection of users to other pages (DG2).*

An extraneous load can also be reduced when the redundancy effect is mediated (Sweller et al., 1998). A redundancy effect occurs when multiple sources of information are made to be self-contained and can be used without referencing each other. Frequently, the same material is presented on multiple occasions in slightly different forms. If subjects find it difficult to ignore the presence of connected materials that they need not process, their cognitive load is increased. Subsequently, redundant and unnecessary duplications of essential material need to be eliminated to reduce the extraneous cognitive load. Similarly, nonessential content and related subject actions (e.g., means for interface adjustment) along with related decorative graphics should be removed (Thüring et al., 1995). Thus, we propose that designers should avoid the redundancy effect: a certification presentation should *mediate the redundancy effect by removing redundant and nonessential content, graphics, and subject actions (DG3).*

Finally, *"germane"* cognitive load is defined "as the cognitive load required to learn, which refers to the working memory resources that are devoted to dealing with intrinsic cognitive load rather than extraneous cognitive load" (Sweller et al., 2019, p. 264).[3] Sweller et al. (2019) assume that rather than contributing to the total cognitive load, germane load redistributes working memory resources from extraneous activities to activities directly relevant to learning by dealing with information intrinsic to the task. The more resources that must be devoted to dealing with extraneous cognitive load, the less free working memory will be available for germane load, ultimately reducing individuals' abilities to deal with intrinsic cognitive load and so less will be learned. In contrast to intrinsic and extraneous load that highly depend on the characteristics of the information (i.e., element interactivity), germane load is independent of the information presented and instead only concerned with characteristics of the subject (i.e., how much working memory resources the subject devotes to deal with the intrinsic load) (Sweller, 2010). We, therefore, do not propose design guidelines for germane load but rather recommend designing certification presentations in a way to manage intrinsic and reduce extraneous load to allow users allocating working memory for germane load, thereby maximizing learning.

Design Guidelines to Enhance Argument Quality

Arguments typically consist of six interrelated elements, regardless of setting (Toulmin, 1958). We next focus on three of the elements strongly enhancing argument quality, while the remaining three elements are commonly left unexpressed as they modify or offer conditions for the validity of the argument (D. Kim & Benbasat, 2006; VerLinden, 1998). First, we address the basic *claim*: an assertion or conclusion proposed for acceptance. Such arguments can also have controversial content. Second, we examine *data*: statements specifying the facts, evidence or previously established beliefs about a situation supporting the claim. Third, we include the *backing*: evidence and related inference such as statistics or other inferences warranting why claims and data should be connected and the claim should be accepted. In general, users accept an argument's claim, if they accept the data and are convinced of their veracity. If a user expresses skepticism about data items, then backing provides reasons why the data should be accepted. In line with this reasoning, including claim-only arguments exhibits lower argument quality since the strength of an argument decreases with a lack of or with deficient evidence (Smith et al., 1991). Some studies suggest that trust-assuring arguments are likely to increase user trust when they include a claim plus data and backing (D. Kim & Benbasat, 2006). Consequently, *assurance statements will have higher argument quality, if they include the claim, data, and backing (DG4).*

Findings confirmed that assurance statements comprising of at least a claim and data will also increase trustworthiness (D. Kim & Benbasat, 2006). Depending on the importance of the statement and the length of the assurance text, designers may decide whether to forego backing to reduce the amount of information needing to be processed. Overall, *assurance statements will enhance argument quality, if they include at least a claim and data (DG5).*

Development and Evaluation Phase: Validating Certification Design Guidelines

Design Propositions

To validate whether the proposed design guidelines will satisfy the meta-requirements, we derived a set of testable design propositions (TDPs) (Arazy et al., 2010; Walls et al., 1992). A design science scholar can articulate numerous testable design propositions to address the extent to which proposed design guidelines satisfy meta-requirements. However, within the context of a single study, only a few propositions can be articulated and tested (Arazy et al., 2010; Walls et al., 1992). For this study, we articulated three propositions (Table 7). First, to test whether our design guidelines can reduce the cognitive load (TDP1), we combine DG1, DG2, and DG3. Second, we use DG4 to increase argument quality (TDP2). Finally, we validate a certification presentation design expected to be highly effective since it imposes a lower cognitive load and improves argument quality (TDP3).

Insert Table 7 About Here

Testing Design Propositions

We tested the three TDPs by using a between-subject design. See Appendix C for detailed information about the experiment, including details on the experiment design and corresponding screenshots, a pre-test as manipulation check, the measurement, and the data gathering and data analyses.

Experimental Design. We developed four instantiations of our design artifact (i.e., certifications presentations) and randomly assigned subjects to test the propositions in a 2x2 between-subjects experiment: (not) implementing DG1-DG3 and (not) implementing DG4. First, we applied no design guidelines for the control group *(none)* and therefore designed the certification presentation as a single block of text without paragraphs or text structuring, thereby imposing a high cognitive load and has low argument quality. Second, we designed a certification presentation *(CP1)* that reduces cognitive load by applying DG1 (i.e., segmenting the content), DG2 (i.e., hierarchical structure and hovering windows for additional information), and DG3 (i.e., removed redundant and nonessential content). Third, we designed a certification presentation *(CP2)* enhancing argument quality by adding assurance arguments including claims, data, and backing (in line with DG4). Finally, we designed a certification presentation *(CP3)* by manipulating *both* the cognitive load and argument quality, applying the two previous treatments simultaneously. We pre-tested the experiment design with 146 subjects from Amazon Mechanical Turk, proving that the certification presentations have the effects expected by the design guidelines.

Data Gathering and Measurement. We next recruited 400 subjects from Amazon Mechanical Turk and removed 28 responses from the final sample, resulting in 372 usable responses that fulfilled our data quality requirements. The subjects participating in the experiment were evenly distributed by gender (female 48%); were, on average, 37 years old (minimum 18 years, maximum 68 years); and had a high school (12%), college (25%), undergraduate (50%) or postgraduate degree (10%). Further, the subjects were employed (73%), looking for work (8%), students (6%), not looking for work (6%), or retired or unable to work (5%). The measurement instrument used the same scales as we applied while validating the theoretical model.

Statistical Analysis. Analyses validated internal consistency, convergent validity, discriminant validity, and CMB is not a major issue in our data. All treatment groups were structurally equivalent for the control variables. Next, we ran a MANOVA to test whether the treatments significantly differed regarding cognitive load and argument quality. The one-way MANOVA showed a statistically significant difference between the treatments on the combined dependent variables ($F(6, 734) = 11.794$, $p < .001$). Post-hoc univariate ANOVAs were conducted, revealing a statistically significant difference between the treatments for *cognitive load* ($F(3, 368) = 21.590$, $p < .001$) and *argument quality* ($F(3, 368) = 5.38$, $p < .001$). Figure 5 depicts that implementing CP1 reduces the *cognitive load*. In contrast, implementing CP2 does not increase *argument quality* but increases the *cognitive load*. Such a design guideline thus has an adverse effect on increasing the cognitive effort needed to process the certification presentation. In addition, we observe that *argument quality* increases beyond the effect of using CP1 singularly if both design propositions are implemented simultaneously. Similarly, the additional cognitive load of implementing only CP2 is reduced when both design propositions are implemented, resulting in a cognitive load equal to that of implementing CP1 alone.

Insert Figure 5 About Here

We then performed post-hoc tests for each dependent variable to check whether these differences are significant (Appendix C). LSD post-hoc analysis on cognitive load between the CP1 and the other treatments revealed a significant difference between CP1 and CP2, $p < .001$ ($M_{Diff} = -1.080$, 95%-CI[-1.386, -.774]), and a weak significant difference between the control group and CP1, $p = .053$ ($M_{Diff} = .301$, 95%-CI[-.004, .606]), **supporting proposition TDP1**. Notably, implementing DG4 in CP2 increase users' cognitive load significantly, as the comparison between the control group and CP2, $p < .001$ ($M_{Diff} = -.779$, 95%-CI[-1.086, -.473]), and between CP1 and CP2 reveals. LSD post-hoc analysis on argument quality revealed no significant difference between CP2 and CP1 or the control group, **not supporting proposition TDP2**.

Between CP3 and the other treatments, we identified a significant difference between the control group and CP3 on cognitive load, p = .046 (M_{Diff} = .312, 95%-CI[.006, .618]); CP2 and CP3 on cognitive load, p < .001 (M_{Diff} = 1.091, 95%-CI[.784, 1.398]); the control group and CP3 on argument quality, p = .005 (M_{Diff} = -.562, 95%-CI[-.954, -.170]); CP1 and CP3 on argument quality, p = .013 (M_{Diff} = -.497, 95%-CI[-.888, -.105]); and CP2 and CP3 on argument quality, p < .001 (M_{Diff} = -.775, 95%-CI[-1.169, -.382]). Thus, CP3 significantly reduced cognitive load and increased argument quality, **supporting proposition TDP3**.

To delve deeper into the relationship between our design instantiations and imposed cognitive load, we performed an additional group comparison on the different cognitive load types (Sweller et al., 1998).[4] We split our cognitive load measurements and assigned the items to each cognitive load time: intrinsic ("*how much mental and perceptual activity was required*"), extraneous ("*how much navigational activity was required*"), and germane ("*how hard did you have to work to accomplish the given task*"), as proposed by Gerjets et al. (2004) and similar to Leppink et al. (2013).

The one-way MANOVA showed a statistically significant difference between the treatments on the combined dependent variables (F(9, 890.898) = 12.754, p < .001). We therefore performed further post-hoc tests for each cognitive load type. Figure 6 outlines the findings and Appendix C provides details for post-hoc test results.

Insert Figure 6 About Here

Performing LSD post-hoc tests reveals that CP1 imposes a lower intrinsic (M_{Diff} = -.90, p < .001), higher extraneous (M_{Diff} = .62, p = .004), and lower germane load than the control group (M_{Diff} = -.56, p = .023). CP3 has similar significant differences to the control group (Appendix C). These findings further substantiate TPD1 because CP1 and CP3 implemented DG1 that manages the intrinsic load. The increment of extraneous is also reasonable because implementing DG1 and DG2 increased users' navigational demands, such as scrolling or clicking on hovering windows, compared to the control group and CP2 treatment. Besides, CLT provides reasons why we witness a decrease in germane load in the treatments CP1 and CP3. If extraneous cognitive load is increased, germane cognitive load is reduced because the user is using working memory resources to deal with the extraneous elements imposed by the design rather than the essential, intrinsic information (Sweller, 2010). Post-hoc tests thereby revealed a boundary condition of DG1 and DG2: if users must devote too much working memory resources to deal with navigational demands (i.e., extraneous load), fewer resources will be available to deal with intrinsic cognitive load, reducing learning in worst cases (Sweller et al., 2019).

In regard to CP2, results highlight that CP2 imposes a higher intrinsic (M_{Diff} = .88, p < .001), an equal extraneous (M_{Diff} = .29, p = .186), and a higher germane load than the control group (M_{Diff} = 1.21, p < .001). If intrinsic cognitive load is high and extraneous low, germane cognitive load will be high because the learner must devote a large proportion of working memory resources to dealing with the essential learning materials (Sweller, 2010). The prevalent high intrinsic and high germane load may provide an indication that users had to put a lot of mental resources in the task of informing themselves about the certification, comprising complex arguments.[5]

Discussion on Design Propositions

The results show that applying our design guidelines DG1, DG2 and DG3 reduces the *cognitive load*. It confirms the effect of design proposition TDP1: *a certification presentation that employs segmentation, organizes the content hierarchically, presents additional information in a hovering window, and removes redundancies imposes a lower cognitive load*. However, implementing DG4 solely to enhance *argument quality* does not directly lead to the expected results (i.e., it does not independently improve perceived argument quality), negating TDP2. Our results indicate an important (and expected) finding regarding the relationship between the *cognitive load* and *argument quality*. Increasing *argument quality* (i.e., more interrelated information elements) increases users' *cognitive load.* Hence, manipulating *argument quality* alone will increase information overload, and users will not be able to assess properly the quality of the assurance information. Simultaneous manipulations of CP3 (i.e., improve argument quality and simultaneously reduce the cognitive load) suggest that not only is the *cognitive load* reduced but also *argument quality* is increased. Consequently, our study suggests that certification presentations must be designed so they reduce the *cognitive load* and enhance *argument quality* to achieve the expected benefits of increasing assurance levels and improving users' trust perceptions.

Conclusion Phase: Discussing Findings on Certification Design

Principal Findings

Our findings suggest that providing a seal on system presentations does not have any significant effect on a user's *perceived assurance* and *trust perceptions*. To restore the expected effects of certification presentation, we show that system owners and certification authorities need to enrich current certification presentations with detailed, structured, and verifiable assurance information. We propose several guidelines for designing certification presentation and organizing assurance information founded on select reference theories. Our findings emphasize that certification presentation designers need to simultaneously seek to reduce the cognitive load of processing assurance statements while enhancing the argument quality of those statements.

With this study, we have applied a thorough theory-driven design approach, resulting in a design proposal, an applied theoretical model, meta-requirements, design guidelines, prototypical instantiations of IS certifications, and validated testable design propositions. We conclude this study by synthesizing these design components into a design theory for certification presentations (Table 8). In general, an "an IS design theory shows the principles inherent in the design of an IS artifact that accomplishes some end, based on knowledge of both IT and human behaviour" (Gregor & Jones, 2007, p. 322). A design theory commonly includes eight core components: (1) the purpose and scope, (2) the constructs, (3) the principles of form and function, (4) the artifact mutability, (5) testable propositions, (6) justificatory knowledge, (7) principles of implementation, and (8) expository instantiation. With our comprehensive and multiple-method design approach, we were able to provide initial descriptions for each component, paving the way for future research on IS certification presentations.

Insert Table 8 About Here

Contribution to the Research

This study's major contribution to the research is threefold (Table 9). First, with this study we contribute to solving the prevalent paradox of certification effectiveness research (Löbbers et al., 2020). While research has frequently analyzed whether or not the presence of a certification seal will change users' *perceived assurance* and *trust perceptions* (e.g., McKnight et al., 2004; Özpolat et al., 2013), prior research only analyzes the impact of certifications from a black box perspective (i.e., embedded vs. not embedded a seal), and lacks a deep understanding about how users actually process certifications (Lansing et al., 2018; Lynn et al., 2016). By building on the ELM, we open the black box of users' information processing and show that observed inconsistencies concerning certification effectiveness likely stem from the shortcoming of using mere graphical cues as certification presentations. Our results highlight that certification presentations should facilitate users' peripheral processing by comprising a graphical seal and facilitate users' cognitive processing by providing detailed assurance information that richly informs users of underlying assurances. We thereby not only answer frequent calls of researchers to improve the design of certification presentations (cf. D. J. Kim, Steinfield, & Lai, 2008; Kimery & McCord, 2006; Lansing et al., 2018; Lowry et al., 2012), but also help users to process and understand the nature and role of certifications and underlying assurances, leading to expected certification effects (Lowry et al., 2012).

Insert Table 9 About Here

While it is reasonable to assume that providing additional assurances impacts users' perceptions in general, providing detailed third-party information about certification' assurances gains high importance because it counteracts users' limited understanding and prevents harmful expectation gaps. This phenomenon holds particularly in electronic markets in which detailed assurance information is required to change users' behaviors (e.g., in cloud service markets). It is also surprising that a certification as valuable and costly third-party information is commonly reduced to mere graphical seals, whereas, in contrast, system owners typically embed comprehensive security and privacy policies as first-party information or show numerous (and detailed) user reports about system usefulness as second-party information. We therefore call for more research on effective presentation of third-party information besides seals, given its unique value in contrast to first- and second-party information, such as being independent, well-founded and verifiable (Löbbers et al., 2020; Özpolat et al., 2013).

Our results should not be viewed in isolation, but synthesized with prior research findings on system certifications. Lansing et al. (2018) stress that certifications' content and codification style (defined as the style of the formal specification in which content dimensions are codified) are important characteristics impacting users' decisions on whether or not to rely on certifications. Likewise, prior research highlights that certifications must include the most appropriate set of assurances for their respective target groups to reduce respective uncertainties (Lansing et al., 2019), given that single assurances differ in their importance to users (Hu et al., 2010; Lansing et al., 2019). Our related research highlights certification presentations can be enhanced by embedding peripheral cues (i.e., authority, social proof and liking cues) to achieve certifications' effects if users are unmotivated or incapable to read underlying assurances (Lins & Sunyaev, 2022). This study's results thus complement prior research because it provides efficient means for providing persuasive assurance information to (potential) users.

Second, while prior research has focused on analyzing just the design of seals (e.g., Aiken & Boush, 2006; D. J. Kim, Steinfield, & Lai, 2008; Moores, 2005), we are the first in providing novel design knowledge about certification presentations in the form of a set of meta-requirements and design guidelines that overcome users' limited understanding of certifications and related assurance information. By building on CLT and Toulmin's model of argumentation as kernel theories, we uncover that the cognitive load and argument quality expressed by certification presentations become two major boundary conditions of certification effectiveness. In particular, we show that an increased cognitive load will negatively influence, and higher argument quality will positively influence users' *perceived assurance* and *trust perceptions*. The identification of cognitive load and argument quality helps to break up prior research's black box perspective on users' processing of certification presentations and provides the foundation for deriving appropriate design guidelines. When elaborating on these guidelines, they appear to be intuitive but also provide advice to efficiently handle the cognitive load and increase argument quality of certification presentations. It is surprising that neither prior research on certification (e.g., Hu et al., 2010), nor practice have considered or came up with related guidelines. For example, the certification *"Certified Privacy"* for webshops, provides information in the form of a comprehensive criteria catalog, which can be downloaded and comprises dozens of complex assurance pages.

By anchoring these insights into the design of certification presentations, we derive several components to formulate a design theory (Gregor & Jones, 2007) for impactful certification presentations (Table 8) and provide an explanation of why the design works. This design theory can be regarded as an improvement (Gregor & Hevner, 2013), because it explains how the problem of limited certification effectiveness can be solved more efficiently by providing assurance information that is properly designed (Hevner et al., 2004).

Third, we contribute to research on CLT and Toulmin's model of argumentation. Whereas CLT aims to develop instructional design guidelines in complex learning tasks (i.e., students learning algebraic problems or training health professionals) (Sweller, 1994; Van Merriënboer & Sweller, 2010), we applied principles of CLT in the context of design guidelines for improving certification presentations. We not only show that a high cognitive load hampers understanding of users when reading online assurance information and thereby decreases the effectiveness of assurance mechanisms. With this exaptation (Gregor & Hevner, 2013) of CLT principles, we also reveal that effective certification presentations can control high element interactivity as well as split attention and redundancy effects while applying derived design guidelines.

In regard to Toulmin's model of argumentation, our research provides new insights into the interdependencies between argument structuring and the cognitive load. Taking a close look at prior research that has applied Toulmin's model of argumentation reveals that this stream has overlooked the downside of applying a complex argument structure. This challenge came apparent in our research. Researchers have frequently applied Toulmin's model of argumentation in online contexts, as recommended (e.g., by D. Kim & Benbasat, 2010), but mostly modified a single argument as shown (e.g., Racherla et al., 2012; Riedl et al., 2010; Yi et al., 2013). For example, in D. Kim and Benbasat (2006)'s experiment subjects viewed each argument (consisting of claim, data, and backing) as an individual piece, independent from the other claims. Our study complements their findings by revealing the positive effect of increased argument quality on the cognitive load–an effect we detected in certification treatment CP2 (i.e., enhancing only argument quality). By comparing the design CP2 with CP3, we revealed a boundary condition regarding the effectiveness of applying Toulmin's model of argumentation when having a need to provide several interrelated assurance arguments. Our findings emphasize that assurance arguments need to include a claim, data, and backing to enhance argument quality, but the cognitive load should be reduced simultaneously by organizing the statements per the design guidelines suggested above. We recommend researchers refraining to use Toulmin's model of argumentation as a silver bullet and consider potential negative side effects of complex argument structures unless they are organized in ways that reduces cognitive load.

Contribution to the Practice

Our study findings also have implications for system owners, certification authorities and users. For system owners, we raise red flags regarding the ongoing simplification of certification deployment in the context of IS use to a mere representation of graphical seals. These seals fail to convey the necessary assurance information effectively. The results show that merely embedding a graphical seal into a system presentation (such as a website) does not lead to the certification's intended effects (i.e., increasing users' perceived assurance and trust perceptions). However, system owners can seek to convey reliable assurance information to overcome the current shortcomings of mere graphical seals. Therefore, system owners should provide detailed assurance information by providing users with informative, transparent and accurate assurance information embedded in properly designed certification presentations.

For certification authorities, our study reveals the initial reasoning for users' limited understanding of the certification. A recent survey shows that only 19% of respondents knew what certifications for webshops were (Statista Survey, 2017). With our study, we highlight that representing certifications as mere graphical cues is insufficient. Instead, certification authorities need to provide detailed assurance information about the provided assurances, the certification process, and the certification authorities' qualities, among other things. To support certification authorities, we derived and validated guidelines for designing certification presentations that improve their impact. Certification presentation designers should focus on enhancing cognitive processing to increase the impact of their certifications, particularly by managing the intrinsic and reducing the extraneous cognitive load while organizing assurance statements using a well-established argument structuring method.

With our study, we also want to motivate users to inform themselves about issued certifications, in particular, to assess carefully the assurances a certification provides because certifications typically differ in their content (Lansing et al., 2019). Users can only make well-informed decisions and select an appropriate system for their use when they are aware of the assurances a certification provides. Finally, exploiting the certification trend, malicious system owners also started to embed self-made assurance seals (so-called *"fake seals")* on their websites (Balboni & Dragan, 2018). Thus, informing oneself about the certification's objectives and its credibility before deciding whether to depend on the (unknown) certification is even more crucial today for users.

Limitations and Future Research

We run our experiments in the context of cloud systems because currently such systems are characterized by a high degree of uncertainty, risk, and ambiguity. We specifically chose storage systems as an empirical setting because they share the typical uncertainties (i.e., considering security and privacy issues). Consequently, our research findings are applicable both to cloud systems and to other online systems with similar characteristics and uncertainties. However, the generalizability and transferability of our findings to general IS system users might be limited because such users might not have a pressing need to inform themselves about system owners.

Our experiments may remain limited in terms of treatment variety and scope. We did not consider alternative or different types of seals in our treatment groups because we focused on differences between the effects of a single seal and the combined effects of a seal and assurance information. Our experiments do not entirely replicate an authentic real-life scenario given that we used a hypothetical cloud system, along with a fictional seal and certification. Nevertheless, the subjects indicated that they believed the given scenario was highly realistic (mean rating of six on a seven-point Likert-type scale). We also conducted our study in the single context of cloud system, given their high practical relevance. Future research should validate our findings in different contexts, such as webshops. To realize the cloud system scenario, we chose a design with a similar visual appearance to that of a real cloud storage system but based on our own arrangement and adjustments (e.g., including an adapted website layout and a fictional logo). Measures of our dependent variables (i.e., users' *trust perceptions*) may differ under real operating conditions. Furthermore, our fictional, experimental implementation of a certification presentation may also affect the validity of testing the design propositions (e.g., different assurance statements may result in either a more positive or a more negative effect). Therefore, the certification presentation used in our experiments serves as a typical example and not as a one-to-one implementation instruction for practical purposes. Moreover, we tested only three TDP and implemented four instantiations of our design artifact since a full factorial experimental design that would consider each design guideline separately would exceed the scope of our primary study. In combination, we would need to test 16 (= 4^2) different treatments. Our evaluation approach thus rather allows for investigating the overall usefulness of derived design guidelines but not for investigating the usefulness of each single design guideline.

Future research should focus on visual aspects (e.g., positioning) to properly display assurance information. In our experiments, we simulated users implicitly clicking on the seal to retrieve further information. In line with previous research that called for improvements in users' click-through rates on seals (D. J. Kim et al., 2017; Kovar et al., 2000; Moores, 2005; Odom et al., 2002), we detected that only 9% to 13% of study subjects (number varied depending on different treatments) clicked the seal on their own initiative without being primed to do so. Consequently, to ensure the impact of certification representation, we recommend that system owners integrate (parts of) the assurance information directly onto their IS presentations, for example, as side information on the main website.

Conclusion

Certifications are now used widely to signal assurances and the institution-based trustworthiness of related system operations. In this regard, they reduce uncertainties and "friction" in commercial exchanges and other interactions between users and system owners. System owners submit to comprehensive and expensive attestations to comply with certification requirements to achieve such effects. They proudly present assurance seals granted through such attestations on their websites or system interfaces. However, the limited effectiveness of using assurance seals as certification presentations to increase assurance or improve trust perceptions calls into question the value of the system owners' efforts. Users are often unfamiliar with seals; they tend to have a limited understanding of underlying assurances or simply misunderstand them. This study questions the current practice of solely embedding graphical seals as a certification presentation and examines whether better certification presentations can be designed.

We adopt a DSR approach and deploy a four-phase research design to clarify how to design impactful IS certification presentations. We not only identify sources of users' limited understanding of seals and formulate a proposal for a seal-based certification presentation augmented with richer assurance information by drawing upon the elaboration likelihood model. We also formulate and validate a set of meta-requirements and design guidelines to improve certification presentations, using CLT and Toulmin's model of argumentation as kernel theories. We developed and evaluated three new certification presentations to validate these guidelines, and thereby revealed that certification presentations increase users' assurance and trust perceptions when the presentations align with users' cognitive information processing needs in ways that reduce their cognitive load and enhance argument quality.

Notes

[1] Note that this breadth of research has led to varying terminology (e.g., web seals, assurance services, certifications). We follow recent conceptualizations of certifications in the IS discipline (e.g., Lansing et al., 2018; Lins et al., 2020; Löbbers & Benlian, 2019) that attest qualities of IS (e.g., security and data protection) and related management practices.
[2] Furthermore, we also note that some of our latent variables have quite non-normal distributions, which is also an often-cited reason for using PLS-SEM although there are also CB-SEM estimators that can cope with non-normal data, but often fail to converge in complex models (Hair et al., 2019).
[3] Please note that germane cognitive load has been specified in differential manner, leading to ambiguous definitions, and resulting in re-conceptualizations over time by Sweller (2010) and Sweller et al. (2019).
[4] We thank the anonymous reviewers for this advice.
[5] We acknowledge that these post-hoc findings should be interpreted with caution because there is an ongoing debate on whether measuring different types of cognitive load is actually valid (refer to Leppink et al., 2013; Sweller et al., 2019, among others).

References

Adam, M., Niehage, L., Lins, S., Benlian, A., & Sunyaev, A. (2020). Stumbling over the Trust Tipping Point: The Effectiveness of Web Seals at Different Levels of Website Trustworthiness. In *Proceedings of the 28th European Conference on Information Systems (ECIS),* An Online AIS Conference.

Aiken, K. D., & Boush, D. M. (2006). Trustmarks, Objective-Source Ratings, and Implied Investments in Advertising: Investigating Online Trust and the Context-Specific Nature of Internet Signals. *Journal of the Academy of Marketing Science*, *34*(3), 308–323. https://doi.org/10.1177/0092070304271004

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, *84*(3), 488–500. https://doi.org/10.2307/1879431

Angst, C. M., & Agarwal, R. (2009). Adoption of Electronic Health Records in the Presence of Privacy Concerns: The Elaboration Likelihood Model and Individual Persuasion. *MIS Quarterly*, *33*(2), 339–370.

Arazy, O., Kumar, N., & Shapira, B. (2010). A Theory-Driven Design Framework for Social Recommender Systems. *Journal of the Association for Information Systems*, *11*(9), 455–490.

Archimi, C. S., Reynaud, E., Yasin, H. M., & Bhatti, Z. A. (2018). How Perceived Corporate Social Responsibility Affects Employee Cynicism: The Mediating Role of Organizational Trust. *Journal of Business Ethics*, *151*(4), 907–921. https://doi.org/10.1007/s10551-018-3882-6
Areni, C. S., & Lutz, R. J. (1988). The Role of Argument Quality in the Elaboration Likelihood Model. *Advances in Consumer Research*(15), 197–203.
Baddeley, A. (2003). Working Memory and Language: An Overview. *Journal of Communication Disorders*, *36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4
Balboni, P., & Dragan, T. (2018). Controversies and Challenges of Trustmarks: Lessons for Privacy and Data Protection Seals. In R. Rodrigues & V. Papakonstantinou (Eds.), *Information Technology and Law Series. Privacy and Data Protection Seals* (Vol. 28, pp. 83–111). T.M.C. Asser Press. https://doi.org/10.1007/978-94-6265-228-6_6
Baskerville, R. (2008). What design science is not. *European Journal of Information Systems*, *17*(5), 441–443. https://doi.org/10.1057/ejis.2008.45
Baskerville, R., Kaul, M., & Storey, V. C. (2018). Aesthetics in design science research. *European Journal of Information Systems*, *27*(2), 140–153. https://doi.org/10.1080/0960085X.2017.1395545
Behrend, T., Sharek, D., Meade, A., & Wiebe, E. (2011). The Viability of Crowdsourcing for Survey Research. *Behavior Research Methods*, *43*(3), 800–813. https://doi.org/10.3758/s13428-011-0081-0
Benlian, A., Kettinger, W. J., Sunyaev, A., & Winkler, T. J. (2018). Special Section: The Transformative Value of Cloud Computing: A Decoupling, Platformization, and Recombination Theoretical Framework. *Journal of Management Information Systems*, *35*(3), 719–739. https://doi.org/10.1080/07421222.2018.1481634
Bhattacherjee, A., & Sanford, C. (2006). Influence Processes for Information Technology Acceptance: An Elaboration Likelihood Model. *MIS Quarterly*, *30*(4), 805–825.
Britt, M., & Gabrys, G. (2002). Implications of Document-level Literacy Skills for Web Site Design. *Behavior Research Methods, Instruments, & Computers*, *34*(2), 170–176. https://doi.org/10.3758/BF03195439
Buchanan, T., Paine, C., Joinson, A. N., & Reips, U.-D. (2007). Development of Measures of Online Privacy Concern and Protection for Use on the Internet. *Journal of the American Society for Information Science and Technology*, *58*(2), 157–165. https://doi.org/10.1002/asi.20459
Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in Problem Solving. In R. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence* (pp. 7–76). Erlbaum.
Danylak, P., Greulich, M., Lins, S., & Sunyaev, A. (2022). Toward a Unified Framework for Information Systems Certification Internalization. In *Proceedings of the 24th IEEE International Conference on Business Informatics (CBI),* Amsterdam, The Netherlands.
DeStefano, D., & LeFevre, J.-A. (2007). Cognitive Load in Hypertext Reading: A Review. *Computers in Human Behavior*, *23*(3), 1616–1641. https://doi.org/10.1016/j.chb.2005.08.012
DiMaggio, P. (1997). Culture and Cognition. *Annual Review of Sociology*, *23*(1), 263–287. https://doi.org/10.1146/annurev.soc.23.1.263
Eagly, A. H., & Chaiken, S. (1993). *The Psychology of Attitudes*. Harcourt Brace Jovanovich College Publishers.
Evans, J. S. B. T. (1984). Heuristic and Analytic Processes in Reasoning. *British Journal of Psychology*, *75*(4), 451–468. https://doi.org/10.1111/j.2044-8295.1984.tb01915.x
Fisher, R., & Chu, Z. S. (2009). Initial online trust formation: the role of company location and web assurance. *Managerial Auditing Journal*, *24*(6), 542–563. https://doi.org/10.1108/02686900910966521
Flynn, L. R., & Goldsmith, R. E. (1999). A Short, Reliable Measure of Subjective Knowledge. *Journal of Business Research*, *46*(1), 57–66. https://doi.org/10.1016/S0148-2963(98)00057-5
Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.2307/3151312
Gaskin, J. (2016). *Common Method Bias (CMB)*. http://statwiki.kolobkreations.com/index.php?title=Confirmatory_Factor_Analysis#Common_Method_Bias_.28CMB.29
Gaskin, J., & Lim, J. (2017). *CFA Tool*. http://statwiki.kolobkreations.com
Gefen, D. (2000). E-Commerce: The Role of Familiarity and Trust. *Omega*, *28*(6), 725–737.
Gerjets, P., Scheiter, K., & Catrambone, R. (2004). Can Learning from Molar and Modular Worked Examples be Enhanced by Providing Instructional Explanations and Prompting Self-Explanations? *Learning and Instruction*, *16*(2), 104–121.
Gregor, S., & Hevner, A. R. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly*, *37*(2), 337–356.
Gregor, S., & Jones, D. (2007). The Anatomy of a Design Theory. *Journal of the Association for Information Systems*, *8*(5), 312–335.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (Second edition). SAGE.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology: Human Mental Workload* (Volume 52, pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Henseler, J., Ringle, C. M., & Sarstedt, M. (2014). A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105.

Hoffmann, C. P., Lutz, C., & Meckel, M. (2015). Digital Natives or Digital Immigrants? The Impact of User Characteristics on Online Trust. *Journal of Management Information Systems*, *31*(3), 138–171. https://doi.org/10.1080/07421222.2014.995538

Houston, R. W., & Taylor, G. K. (1999). Consumer Perceptions of CPA WebTrustSM Assurances: Evidence of an Expectation Gap. *International Journal of Auditing*, *3*(2), 89–105. https://doi.org/10.1111/1099-1123.00051

Hsu, C., Lee, J.-N., & Straub, D. W. (2012). Institutional Influences on Information Systems Security Innovations. *Information Systems Research*, *23*(3-part-2), 918–939. https://doi.org/10.1287/isre.1110.0393

Hu, X., Wu, G., Wu, Y., & Zhang, H. (2010). The Effects of Web Assurance Seals on Consumers' Initial Trust in an Online Vendor: A Functional Perspective. *Decision Support Systems*, *48*(2), 407–418. https://doi.org/10.1016/j.dss.2009.10.004

Kim, D. J. (2008). Self-Perception-Based Versus Transference-Based Trust Determinants in Computer-Mediated Transactions: A Cross-Cultural Comparison Study. *Journal of Management Information Systems*, *24*(4), 13–45. https://doi.org/10.2753/MIS0742-1222240401

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A Trust-based Consumer Decision-making Model in Electronic Commerce: The Role of Trust, Perceived Risk, and their Antecedents. *Decision Support Systems*, *44*(2), 544–564. https://doi.org/10.1016/j.dss.2007.07.001

Kim, D. J., Steinfield, C., & Lai, Y.-J. (2008). Revisiting The Role of Web Assurance Seals in Business-to-Consumer Electronic Commerce. *Decision Support Systems*, *44*(4), 1000–1015. https://doi.org/10.1016/j.dss.2007.11.007

Kim, D. J., Yim, M.-S., Sugumaran, V., & Rao, H. R. (2017). Web Assurance Seal Services, Trust and Consumers' Concerns. *European Journal of Information Systems*, *25*(3), 252–273. https://doi.org/10.1057/ejis.2015.16

Kim, D., & Benbasat, I. (2006). The Effects of Trust-Assuring Arguments on Consumer Trust in Internet Stores: Application of Toulmin's Model of Argumentation. *Information Systems Research*, *17*(3), 286–300. https://doi.org/10.1287/isre.1060.0093

Kim, D., & Benbasat, I. (2010). Designs for effective implementation of trust assurances in internet stores. *Commun. ACM*, *53*(2), 121. https://doi.org/10.1145/1646353.1646386

Kim, K., & Kim, J. (2011). Third-party Privacy Certification as an Online Advertising Strategy: An Investigation of the Factors Affecting the Relationship between Third-party Certification and Initial Trust. *Journal of Interactive Marketing*, *25*(3), 145–158. https://doi.org/10.1016/j.intmar.2010.09.003

Kimery, K. M., & McCord, M. (2006). Signals of Trustworthiness in E-Commerce: Consumer Understanding of Third-Party Assurance Seals. *Journal of Electronic Commerce in Organizations*, *4*(4), 52–74.

Kimery, K. M., & McCord, M. (2008). Seals on Retail Web Sites: A Signaling Theory Perspective on Third-Party Assurances. In M. Khosrow-Pour (Ed.), *Web Technologies for Commerce and Services Online* (pp. 111–134). IGI Global. https://doi.org/10.4018/978-1-59904-822-2.ch007

Kock, N. (2015). Common Method Bias in PLS-SEM. *International Journal of E-Collaboration*, *11*(4), 1–10. https://doi.org/10.4018/ijec.2015100101

Kock, N., & Lynn, G. (2012). Lateral Collinearity and Misleading Results in Variance-Based SEM: An Illustration and Recommendations. *Journal of the Association for Information Systems*, *13*(7), 546–580. https://doi.org/10.17705/1jais.00302

Kovar, S. E., Burke, K. G., & Kovar, B. R. (2000). Consumer Responses to the CPA WEBTRUST™ Assurance. *Journal of Information Systems*, *14*(1), 17–35. https://doi.org/10.2308/jis.2000.14.1.17

Kuechler, B., & Vaishnavi, V. (2008). On theory development in design science research: anatomy of a research project. *European Journal of Information Systems*, *17*(5), 489–504. https://doi.org/10.1057/ejis.2008.40

Lala, V., Arnold, V., Sutton, S. G., & Guan, L. (2002). The Impact of Relative Information Quality of E-Commerce Assurance Seals on Internet Purchasing Behavior. *International Journal of Accounting Information Systems*, *3*(4), 237–253. https://doi.org/10.1016/S1467-0895(02)00069-6
Lansing, J., Benlian, A., & Sunyaev, A. (2018). “Unblackboxing” Decision Makers’ Interpretations of IS Certifications in the Context of Cloud Service Certifications. *Journal of the Association for Information Systems*, *19*(11), 1064–1096.
Lansing, J., Siegfried, N., Sunyaev, A., & Benlian, A. (2019). Strategic signaling through cloud service certifications: Comparing the relative importance of certifications’ assurances to companies and consumers. *The Journal of Strategic Information Systems*, *28*(4), 1–23. https://doi.org/10.1016/j.jsis.2019.101579
Laurent, G., & Kapferer, J.-N. (1985). Measuring Consumer Involvement Profiles. *Journal of Marketing Research*, *22*(1), 41–53. https://doi.org/10.2307/3151549
Leppink, J., Paas, F., van der Vleuten, C. P. M., van Gog, T., & van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1
Lins, S., Grochol, P., Schneider, S., & Sunyaev, A. (2016). Dynamic Certification of Cloud Services: Trust, but Verify! *IEEE Security & Privacy*, *14*(2), 66–71. https://doi.org/10.1109/MSP.2016.26
Lins, S., Kromat, T., Löbbers, J., Benlian, A., & Sunyaev, A. (2020). Why Don’t You Join In? A Typology of Information System Certification Adopters. *Decision Sciences*, *forthcoming.* https://doi.org/10.1111/deci.12488
Lins, S., Schneider, S., & Sunyaev, A. (2018). Trust is Good, Control is Better: Creating Secure Clouds by Continuous Auditing. *IEEE Transactions on Cloud Computing*, *6*(3), 890–903. https://doi.org/10.1109/TCC.2016.2522411
Lins, S., Schneider, S., Szefer, J., Ibraheem, S., & Sunyaev, A. (2019). Designing Monitoring Systems for Continuous Certification of Cloud Services: Deriving Meta-requirements and Design Guidelines. *Communications of the Association for Information Systems*, *44*(Paper 25), 406–510. https://doi.org/10.17705/1CAIS.04425
Lins, S., & Sunyaev, A. (2017). Unblackboxing IT Certifications: A Theoretical Model Explaining IT Certification Effectiveness. In *Proceedings of the 38th International Conference on Information Systems (ICIS 2017),* Seoul, South Korea.
Lins, S., & Sunyaev, A. (2022). Advancing the Presentation of IS Certifications: Theory-driven Guidelines for Designing Peripheral Cues to Increase Users’ Trust Perceptions. *Preprint*.
Löbbers, J., & Benlian, A. (2019). The effectiveness of IS certification in E-commerce: does personality matter? *Journal of Decision Systems*, *28*(3), 233–259. https://doi.org/10.1080/12460125.2019.1684867
Löbbers, J., Lins, S., Kromat, T., Benlian, A., & Sunyaev, A. (2020). A multi-perspective lens on web assurance seals: contrasting vendors’ intended and consumers’ perceived effects. *Electronic Commerce Research*, *forthcoming.* https://doi.org/10.1007/s10660-020-09415-2
Lowry, P. B., Moody, G., Vance, A., Jensen, M., Jenkins, J., & Wells, T. (2012). Using an Elaboration Likelihood Approach to Better Understand the Persuasiveness of Website Privacy Assurance Cues for Online Consumers. *Journal of the American Society for Information Science and Technology*, *63*(4), 755–776. https://doi.org/10.1002/asi.21705
Lynn, T., van der Werff, L., Hunt, G., & Healy, P. (2016). Development of a Cloud Trust Label: A Delphi Approach. *Journal of Computer Information Systems*, *56*(3), 185–193. https://doi.org/10.1080/08874417.2016.1153887
Mak, B., & Lyytinen, K. (1997). A model to assess the behavioral impacts of consultative knowledge based systems. *Information Processing & Management*, *33*(4), 539–550. https://doi.org/10.1016/S0306-4573(97)00015-0
Mavlanova, T., Benbunan-Fich, R., & Lang, G. (2016). The Role of External and Internal Signals in E-commerce. *Decision Support Systems*, *87*, 59–68. https://doi.org/10.1016/j.dss.2016.04.009
Mayer, R. E., & Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, *38*(1), 43–52.
McKnight, D. H., Choudhury, V., & Kacmar, C. J. (2002). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, *13*(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81
McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial Trust Formation in New Organizational Relationships. *The Academy of Management Review*, *23*(3), 473–490.

McKnight, D. H., Kacmar, C. J., & Choudhury, V. (2004). Shifting Factors and the Ineffectiveness of Third Party Assurance Seals: A Two-Stage Model of Initial Trust in a Web Business. *Electronic Markets*, *14*(3), 252–266. https://doi.org/10.1080/1019678042000245263

Mongeau, P. A., & Stiff, J. B. (1993). Specifying Causal Relationships in the Elaboration Likelihood Model. *Communication Theory*, *3*(1), 65–72. https://doi.org/10.1111/j.1468-2885.1993.tb00057.x

Moores, T. (2005). Do Consumers Understand the Role of Privacy Seals in e-Commerce? *Communications of the ACM*, *48*(3), 86–91. https://doi.org/10.1145/1047671.1047674

Nunamaker, J. F., Chen, M., & Purdin, T. D. (1990). Systems Development in Information Systems Research. *Journal of Management Information Systems*, *7*(3), 89–106. https://doi.org/10.1080/07421222.1990.11517898

Nunnally, J. C. (1978). *Psychometric Theory* (2d ed). *McGraw-Hill series in psychology*. McGraw-Hill.

Odom, M. D., Kumar, A., & Saunders, L. (2002). Web Assurance Seals: How and Why They Influence Consumers' Decisions. *Journal of Information Systems*, *16*(2), 231–250.

Österle, H., Becker, J., Frank, U., Hess, T., Karagiannis, D., Krcmar, H., Loos, P., Mertens, P., Oberweis, A., & Sinz, E. J. (2011). Memorandum on design-oriented information systems research. *European Journal of Information Systems*, *20*(1), 7–10. https://doi.org/10.1057/ejis.2010.55

Özpolat, K., Gao, G., Jank, W., & Viswanathan, S. (2013). Research Note—The Value of Third-Party Assurance Seals in Online Retailing: An Empirical Investigation. *Information Systems Research*, *24*(4), 1100–1111. https://doi.org/10.1287/isre.2013.0489

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture. *Instructional Science*, *32*(1/2), 1–8. https://doi.org/10.1023/B:TRUC.0000021806.17516.d0

Paas, F., Tuovinen, J. E., Tabbers, H., & van Gerven, P. W. M. (2003). Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist*, *38*(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a Sufficient Condition for Data Quality on Amazon Mechanical Turk. *Behavior Research Methods*, *46*(4), 1023–1031. https://doi.org/10.3758/s13428-013-0434-y

Peffers, K., Tuunanen, T., & Niehaves, B. (2018). Design science research genres: introduction to the special issue on exemplars and criteria for applicable design science research. *European Journal of Information Systems*, *27*(2), 129–139. https://doi.org/10.1080/0960085X.2018.1458066

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Petter, S., Khazanchi, D., & Murphy, J. D. (2010). A design science based evaluation framework for patterns. *ACM SIGMIS Database*, *41*(3), 9–26. https://doi.org/10.1145/1851175.1851177

Petty, R. E., & Cacioppo, J. T. (Eds.). (1986). *Communication and Persuasion*. Springer New York. https://doi.org/10.1007/978-1-4612-4964-1

Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal Involvement as a Determinant of Argument-Based Persuasion. *Journal of Personality and Social Psychology*, *41*(5), 847–855. https://doi.org/10.1037/0022-3514.41.5.847

Petty, R. E., & Wegener, D. (1999). The Elaboration Likelihood Model: Current Status and Controversies. In S. Chaiken & Y. Trope (Eds.), *Dual-Process Theories in Social Psychology.* Guilford Press.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Pollock, E., Chandler, P. A., & Sweller, J. (2002). Assimilating Complex Information. *Learning and Instruction*, *12*(1), 61–86.

Portz, K., Strong, J. M., Busta, B., & Schneider, K. (2000). Do Consumers Understand What WebTrust Means? *CPA Journal*, *70*(10), 46–52.

Racherla, P., Mandviwalla, M., & Connolly, D. J. (2012). Factors affecting consumers' trust in online product reviews. *Journal of Consumer Behaviour*, *11*(2), 94–104. https://doi.org/10.1002/cb.385

Renner, M., Lins, S., & Sunyaev, A. (2021). A Taxonomy of IS Certification's Characteristics. In *ICIEB'21: 2021 2nd International Conference on Internet and E-Business,* Barcelona, Spain.

Riedl, R., Hubert, M., & Kenning, P. (2010). Are There Neural Gender Differences in Online Trust? An fMRI Study on the Perceived Trustworthiness of Ebay Offers. *MIS Quarterly*, *34*(2), 397–428.

Rifon, N., La Rose, R., & Choi, S. M. (2005). Your Privacy Is Sealed: Effects of Web Privacy Seals on Trust and Personal Disclosures. *Journal of Consumer Affairs*, *39*(2), 339–362. https://doi.org/10.1111/j.1745-6606.2005.00018.x
Ringle, C. M., Wende, S., & Becker, J.-M. (2015). *SmartPLS 3* (Version 3) [Computer software]. SmartPLS GmbH, www.smartpls.com. Boenningstedt, Germany. www.smartpls.com
Schneider, S., & Sunyaev, A. (2016). Determinant Factors of Cloud-Sourcing Decisions: Reflecting on the IT Outsourcing Literature in the Era of Cloud Computing. *Journal of Information Technology*, *31*(1), 1–31. https://doi.org/10.1057/jit.2014.25
Shah, M. H., Peikari, H. R., & Yasin, N. M. (2014). The determinants of individuals' perceived e-security: Evidence from Malaysia. *International Journal of Information Management*, *34*(1), 48–57. https://doi.org/10.1016/j.ijinfomgt.2013.10.001
Siegfried, N., Winkler, N., & Benlian, A. (2020). Do bad experiences loom larger than good ones? The role of prior purchase experiences on the effectiveness of IS certifications. *Journal of Decision Systems*, *29*(2), 79–101. https://doi.org/10.1080/12460125.2020.1766235
Simon, H. A. (1996). *The Sciences of the Artificial* (3rd ed.). MIT Press. https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=49230
Singh, R. I., Sumeeth, M., & Miller, J. (2011). A user-centric evaluation of the readability of privacy policies in popular web sites. *Information Systems Frontiers*, *13*(4), 501–514. https://doi.org/10.1007/s10796-010-9228-2
Sloman, S. A. (1996). The Empirical Case for Two Systems of Reasoning. *Psychological Bulletin*, *119*, 3–22.
Smith, G. F., Benson, P. G., & Curley, S. P. (1991). Belief, knowledge, and uncertainty: A cognitive perspective on subjective probability. *Organizational Behavior and Human Decision Processes*, *48*(2), 291–321. https://doi.org/10.1016/0749-5978(91)90016-M
Speier, C., & Morris, M. G. (2003). The Influence of Query Interface Design on Decision-making Performance. *MIS Quarterly*, *27*(3), 397–423.
Statista Survey. (2017). *Do you know what seals of approval for online shops are?* https://www.statista.com/statistics/705448/awareness-about-seal-of-approval-for-online-shopping-in-the-us/
Steelman, Z. R., Hammer, B. I., & Limayem, M. (2014). Data Collection in the Digital Age: Innovative Alternatives to Student Samples. *MIS Quarterly*, *38*(2), 355–378.
Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating Argumentative Knowledge Construction with Computer-supported Collaboration Scripts. *International Journal of Computer-Supported Collaborative Learning*, *2*(4), 421–447. https://doi.org/10.1007/s11412-007-9028-y
Stewart, K. J. (2003). Trust Transfer on the World Wide Web. *Organization Science*, *14*(1), 5–17. https://doi.org/10.1287/orsc.14.1.5.12810
Sturm, B., & Sunyaev, A. (2019). Design Principles for Systematic Search Systems: A Holistic Synthesis of a Rigorous Multi-cycle Design Science Research Journey. *Business & Information Systems Engineering*, *61*(1), 91–111. https://doi.org/10.1007/s12599-018-0569-6
Sunyaev, A., & Schneider, S. (2013). Cloud services certification. *Communications of the ACM*, *56*(2), 33–36. https://doi.org/10.1145/2408776.2408789
Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7
Sweller, J. (1994). Cognitive Load Theory, Learning Difficulty, and Instructional Design. *Learning and Instruction*, *4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5
Sweller, J. (2004). Instructional Design Consequences of an Analogy between Evolution by Natural Selection and Human Cognitive Architecture. *Instructional Science*, *32*(1-2), 9–31. https://doi.org/10.1023/B:TRUC.0000021808.72598.4d
Sweller, J. (2010). Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. *Educational Psychology Review*, *22*(2), 123–138. https://doi.org/10.1007/s10648-010-9128-5
Sweller, J., van Merrienboer, J. G., & Paas, F. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*(3), 251–296. https://doi.org/10.1023/A:1022193728205
Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5
Thüring, M., Hannemann, J., & Haake, J. M. (1995). Hypermedia and Cognition: Designing for Comprehension. *Communications of the ACM*, *38*(8), 57–66. https://doi.org/10.1145/208344.208348
Toulmin, S. E. (1958). *The Use of Argument*. Cambridge University Press.

Trenz, M., Huntgeburth, J., & Veit, D. (2018). Uncertainty in cloud service relationships: Uncovering the differential effect of three social influence processes on potential and current users. *Information & Management*, *55*(8), 971–983. https://doi.org/10.1016/j.im.2018.05.002

Vaishnavi, V., & Kuechler, W. (2015). *Design science research methods and patterns: Innovating information and communication technology* (Second edition). CRC Press LLC. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=935839

Vaishnavi, V., Kuechler, W., & Petter, S. (2017). *Design Science Research in Information Systems*. www.desrist.org/design-research-in-information-systems/

Van Merriënboer, J. J. G., & Sweller, J. (2010). Cognitive Load Theory in Health Professional Education: Design Principles and Strategies. *Medical Education*, *44*(1), 85–93. https://doi.org/10.1111/j.1365-2923.2009.03498.x

van Merriënboer, J. J., & Sweller, J. (2005). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, *17*(2), 147–177. https://doi.org/10.1007/s10648-005-3951-0

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: a Framework for Evaluation in Design Science Research. *European Journal of Information Systems*, *25*(1), 77–89. https://doi.org/10.1057/ejis.2014.36

VerLinden, J. G. (1998). *Verifiers in Everyday Argument: An Expansion of the Toulmin Model*. http://cfile232.uf.daum.net/attach/186A5C4E5022069C2FC796

Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2015). Discriminant Validity Testing in Marketing: An Analysis, Causes for Concern, and Proposed Remedies. *Journal of the Academy of Marketing Science*, *44*(1), 119–134. https://doi.org/10.1007/s11747-015-0455-4

Walls, J. G., Widmeyer, G. R., & El Sawy, O. A. (1992). Building an Information System Design Theory for Vigilant EIS. *Information Systems Research*, *3*(1), 36–59.

Winter, R. (2008). Design science research in Europe. *European Journal of Information Systems*, *17*(5), 470–475. https://doi.org/10.1057/ejis.2008.44

Yang, S.-C., Hung, W.-C., Sung, K., & Farn, C.-K. (2006). Investigating Initial Trust Toward E-Tailers From the Elaboration Likelihood Model Perspective. *Psychology & Marketing*, *23*(5), 429–445.

Ye, L. R., & Johnson, P. E. (1995). The Impact of Explanation Facilities on User Acceptance of Expert Systems Advice. *MIS Quarterly*, *19*(2), 157–172. https://doi.org/10.2307/249686

Yi, M. Y., Yoon, J. J., Davis, J. M., & Lee, T. (2013). Untangling the antecedents of initial trust in Web-based health information: The roles of argument quality, source expertise, and user perceptions of information quality and risk. *Decis. Support Syst.*, *55*(1), 284–295. https://doi.org/10.1016/j.dss.2013.01.029

Zucker, L. G. (1986). Production of Trust: Institutional Sources of Economic Structure, 1840–1920. *Research in Organizational Behavior*, *8*, 53–111.

About the Authors

Sebastian Lins is a postdoctoral researcher and lecturer at the Research Group Critical Information Infrastructures (cii), Institute of Applied Informatics and Formal Description Methods, Karlsruhe Institute of Technology (KIT), Germany. He has received his doctorate degree from the University of Cologne**.** His main interests in the field of information systems research are the (continuous) certification of cloud services and distributed ledger technology as well as understanding and enhancing the effectiveness of IS certifications. His work has been published in international journals such as *IEEE Transactions on Cloud Computing*, *Decision Sciences*, and *Communications of the Association for Information Systems*, as well as in conference proceedings such as *International Conference on Information Systems* and *European Conference on Information Systems*.

Jan-Michael Becker is an associate professor in the department of marketing at the BI Norwegian Business School. He has received his doctorate degree from the University of Cologne**.** His research interest and expertise focus on the digital transformation of marketing and consumer behavior as well as marketing analytics, behavioral research methods and computational statistics. He thereby bridges information systems and marketing topics. His research has been published in several premier academic journals, including *Information Systems Research*, *MIS Quarterly*, *Psychometrika*, *Nature Human Behavior*, *Multivariate Behavioral Research*, *Long Range Planning*, *Journal of Business Research*, and *Marketing Letters*. He is a co-developer of the SmartPLS software application for modeling structural equation models with Partial Least Squares (PLS) path modeling.

Kalle Lyytinen is the Iris S. Wolstein Professor of Management Design at Case Western Reserve University's Weatherhead School of Management, where he is also director of their Doctor of Management program. His research focuses on how organizations design and manage digital innovations. He is notable for his breadth of scholarship and leadership in the field of Information Systems, and has an extensive list of over 300 publications in numerous prestigious journals including *Information Systems Research*, *Management Information Systems Quarterly*, and *Organization Science* and leading conferences. He is currently among the top five scholars in the information system field by citations (H-index 71).

Ali Sunyaev is a professor for computer science at the Karlsruhe Institute of Technology (KIT), Germany. Before joining KIT, he was Professor at the University of Kassel and the University of Cologne. He received his PhD in Information Systems from the Technical University of Munich (TUM). His research work accounts for the multifaceted use contexts of digital technologies with research on human behavior affecting IT applications and vice versa. His research appeared in journals including *Journal of Information Technology*, *Journal of Management Information Systems*, *Journal of the Association for Information Systems*, *IEEE Transactions on Cloud Computing*, *Communications of the ACM*, and others. His research work has been appreciated numerous times and is featured in a variety of media outlets.

Appendix A. Problem Awareness Phase: Why Do Certifications Fail in Achieving Effects

A1. Experiment Procedure and Treatments to Validate Certification Presentation Effects

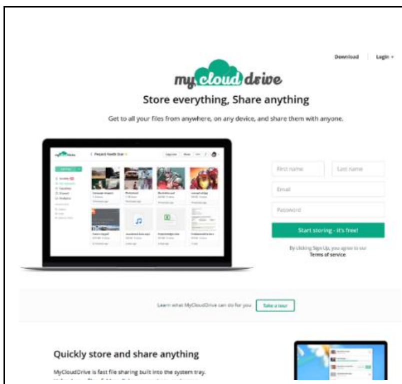
We conducted a between-subject experiment using three groups of subjects to test the hypotheses. We asked subjects to evaluate their usage of a cloud storage system named "*MyCloudDrive*". Please note that *MyCloudDrive* was not introduced as a fictional cloud service; instead, we informed the subjects that it would be released soon to increase the perceived scenario realism.

The experiment was conducted in five steps. First, we provided a short description and examples of current cloud systems. Second, we asked several control questions to control whether potential users' *cloud knowledge, cloud involvement, and disposition to trust* or *privacy concerns* would have a confounding impact on the predicted effects. Third, we introduced a scenario in which the users had to select a cloud storage system for their personal use and in which they would evaluate *MyCloudDrive* as one option. Fourth, we created three between-subject treatments and randomly assigned subjects to one of these treatments: (1) the first treatment group (control group) with *MyCloudDrive*'s website only, (2) the second treatment group with *MyCloudDrive*'s website including a certification presented with only a seal, (3) and the third treatment group with *MyCloudDrive's* website including a certification presentation comprising a seal and assurance information. See Figure A.1. for screenshot snippets of treatments. In the end, we measured users' perceived assurance and trust perceptions, and gathered demographic user information.

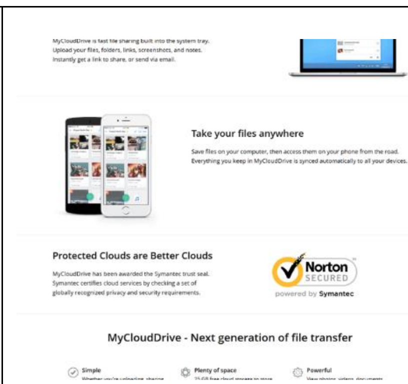
Cloud Service Website. To elicit natural responses from subjects, the website's layout, pictures, and text accurately simulated actual cloud storage provider's websites (i.e., *Dropbox*, *Jumpshare*, and *Infinit*) and were adapted to fit the study objectives. The *MyCloudDrive* website informed subjects about the cloud storage system's main features and characteristics. We aligned storage capacities and prices based on the current storage offers at the time of the study. All subjects had to scroll through the entire website before continuing.

Presentation of the Certification Seal. In our second treatment group, we presented a certification as a seal embedded into the website page. We expected it to trigger users' peripheral processing. In our case, we integrated the *Symantec* seal because pre-tests confirmed it is well known and one of the most impactful seals to be understood by the treatment group. To reduce the experiment's complexity, we did not add other treatment groups that would consider other seals. We expected the most prevalent seal to have the strongest effect on users. In addition, we assumed that if a presentation comprising of a seal and assurance information outperformed the most impactful seal alone, it would outperform other seals as well.

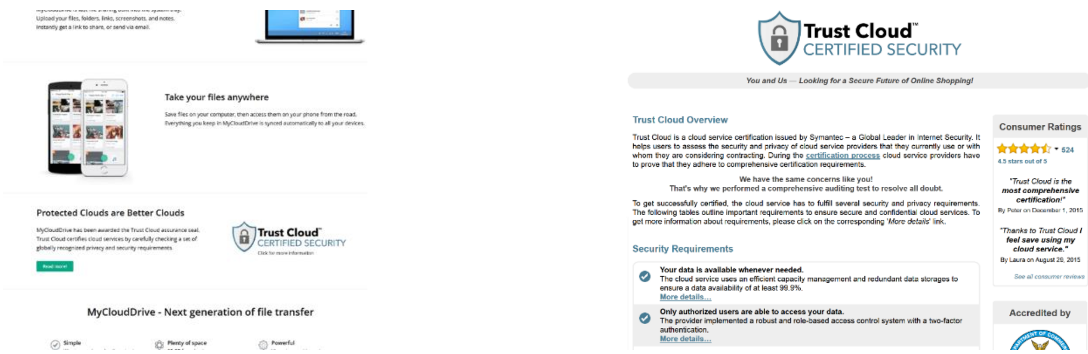
Presentation of the Certification Seal and Assurance Information. For the third treatment group, we used a more complex scenario. We replaced the *Symantec* seal shown on the cloud service's website because the *Symantec* seal does not currently provide required assurance information in the form fitting the purpose of our experiment. Consequently, we designed a fictional certification seal called "*Trust Cloud—Certified Security*", based on some actual seals, to ensure a high level of realism for the experiment and so the seal would fit the study objectives. We also developed a related detailed assurance text that provides subjects detailed information about the certification process related to *Trust Cloud,* including information about the certification authority, the attestation process and the assurance statements, which all would facilitate users' cognitive processing. The assurance information was shown on an additional web page presented immediately after the *MyCloudDrive* website. During the experiment, we also requested the subjects to explore the information on this website.



Snippet treatment website only



Snippet treatment certification seal



Snippet treatment seal and assurance information. Showing the website and the seal related to the certification

Snippet assurance information shown in the seal and assurance information treatment group

Figure A.1. Experiment Design Snippets (Problem Awareness Phase)

A2. Measurement

The experiment used validated scales from the literature for all identified constructs. All items were measured using seven-point Likert-type scales. Because of the fictional nature of the cloud service, the subjects had no experience with or firsthand knowledge about the cloud service provider *MyCloudDrive*, though both experience and firsthand knowledge are required to form general trust (McKnight et al., 1998). Therefore, our study focuses on a period during which a user visits and explores an online system owner's website for the first time, placing the study context within the domain of *initial trust* (McKnight et al., 1998). We adapted measurements for user's *trust perceptions* accordingly. See Table A.1 for an overview of the items used to measure each construct.

Table A.1. Measurement Items

Perceived assurance (D. J. Kim, Ferrin, & Rao, 2008)
I feel confident that MyCloudDrive won't use or sell my personal data for other purposes without my authorization.
I feel confident that MyCloudDrive won't share my personal data with other entities without my authorization.
I feel confident that only authorized persons have access to my personal data stored in MyCloudDrive.
I'm confident of the privacy of my personal data during a transmission (e.g., uploading or downloading files) to MyCloudDrive.
I feel confident that MyCloudDrive implements appropriate security measures to protect stored data.
I feel confident that MyCloudDrive protects my data from accidentally being altered or destroyed during a transmission on the Internet.
I feel save storing my data in MyCloudDrive.
In general, I believe storing data in MyCloudDrive is securer than storing it in other providers' services.
Trust perceptions (McKnight et al., 2002)
I trust that MyCloudDrive keeps my best interests in mind.
I believe that MyCloudDrive will be ready and willing to help me.
I think that MyCloudDrive is interested in my well-being, not just its own.

I would characterize MyCloudDrive as honest.
MyCloudDrive appears to be one who would keep promises and commitments.
MyCloudDrive seems to be trustworthy.
MyCloudDrive seems to be competent and effective in providing their services.
Overall, MyCloudDrive appears to be a capable and proficient cloud service provider.
In general, MyCloudDrive appears to be very knowledgeable about the cloud services.
Cloud service involvement (Laurent & Kapferer, 1985)
I have a strong interest in cloud services.
Cloud services are very important to me.
For me cloud services matter a lot.
Cloud service knowledge (Flynn & Goldsmith, 1999)
Among my circle of friends, I'm one of the "experts" on cloud services.
I do feel very knowledgeable about cloud services.
Compared to most other people, I know much about cloud services.
I know pretty much about cloud services.
Disposition to trust (Gefen, 2000)
I generally have faith in humanity.
I generally trust other people unless they give me reason not to.
I tend to count upon other people.
I feel that people are generally reliable.
Privacy concerns (Buchanan et al., 2007)
In general, how concerned are you about your privacy while you are using the internet?
Are you concerned about online organizations not being who they claim they are?
Are you concerned that you are asked for too much personal information when you register or make online purchases?
Are you concerned about online identity theft?
Are you concerned who might access your stored photos or music files online?
Are you concerned about unknown people obtaining personal information about you from your online activities?
Are you concerned that if you use your credit card to buy something on the internet your credit card number will obtained or intercepted by someone else?
Are you concerned that if you use your credit card to buy something on the internet your card will be mischarged?

A3. Data Gathering

We recruited 300 participants from Amazon Mechanical Turk. Research has demonstrated that the results of surveys using Mechanical Turk respondents have high reliability and provide high-quality data comparable to student samples or online convenience samples (Behrend et al., 2011; Steelman et al., 2014). In addition, Mechanical Turk is a suitable platform for reaching internet-savvy users who have a consummate fit with our experiment setting and goals. We restricted participation to users with a high reputation (at least 99% approval ratings and at least 5,000 conducted tasks), which is sufficient to ensure high data quality (Peer et al., 2014). In addition, we restricted participation in the US to ensure the applied seal (*Symantec*) would be well recognized among all of the participants. Finally, we embedded attention-check questions to the survey and recorded the time spent on each page to remove responses that had received insufficient attention. As a result, we removed seven responses because the subjects failed attention checks or rushed through the experiment (three subjects in the website treatment group, one subject in the seal treatment, and three subjects in the assurance information treatment group). This process resulted in 293 responses fulfilling initial data quality requirements.

A4. Data Analysis

We assessed the construct measurement scales (**Table A.2**). All indicators fulfilled the minimum loading requirements (significance and load value) between the indicator and its corresponding underlying factor, proving convergent validity. The average variance extracted (AVE) was greater than the suggested minimum of .5 (Fornell & Larcker, 1981). Composite reliability (CR) demonstrated good internal consistency of measures in that each construct had CR > .7, exceeding the recommended threshold (Nunnally, 1978). The square root of each construct's AVE exceeded the inter-construct correlations, demonstrating adequate discriminant validity. The heterotrait-monotrait (HTMT) ratio of correlations (Henseler et al., 2014; Voorhees et al., 2015) showed a value less than a threshold value of .85 for all constructs, suggesting no discriminant validity problems.

Table A.2. Construct Reliability and Discriminant Validity

Construct	**CR**	**AVE**	**Inter-construct correlations (HTMT)**					
			(1)	(2)	(3)	(4)	(5)	(6)
(1) Cloud service involvement	.976	.932	.965 (-)					
(2) Cloud service knowledge	.936	.785	.676 (.719)	.886 (-)				
(3) Disposition to trust	.936	.785	.251 (.270)	.207 (.230)	.886 (-)			
(4) Perceived assurance	.963	.767	.268 (.277)	.265 (.275)	.325 (.350)	.876 (-)		
(5) Privacy concerns	.938	.657	-.010 (.037)	-.049 (.075)	-.017 (.057)	-.040 (.060)	.810 (-)	
(6) Trust perceptions	.956	.707	.218 (.216)	.173 (.176)	.358 (.373)	.788 (.823)	.086 (.097)	.841 (-)

Second, both procedural and statistical methods were used to control for common method bias (CMB). Since the data were collected through an online experiment, we applied ex ante the recommendations of Podsakoff et al. (2003) for controlling CMB, including instructing subjects that answers are fully anonymized, that they should take their time to carefully and honestly answer the questions and that no right and wrong answers exist; using a fictional cloud storage system as scenario to prevent bias from subjects' prior opinions (i.e., positive or negative thoughts about existing cloud systems); counterbalancing question orders; using existing, reliable measures; randomizing items; applying (short) temporal (i.e., countdown timer before continuing) and proximal separation (i.e., different pages) of measurements for variables. We conducted a common latent factor (CLF) test using IBM SPSS Amos. We first compared the standardized regression weights of all items for models with and without CLF. The differences in these regression weights were found to be very small (<0.200) (Archimi et al., 2018; Gaskin, 2016). We then conducted a zero constraints test using IBM SPSS Amos using the *'Specific Bias Test'* plugin (Gaskin & Lim, 2017).

The test did not detect significant response bias (χ^2 = 1118,000; df = 956; p = 1). We conclude that a potential CMB does not pose significant threat to our results.

Third, we controlled for whether users' cloud knowledge, cloud involvement, disposition to trust, or privacy concerns differed across treatments to check for the presence of confounding effects and to determine whether the treatment groups were structurally equivalent (Table A.3). We did not find any significant differences for our control variables *cloud knowledge* and *involvement.* However, we did find significant differences in users' *privacy concerns*, which varied across the seal only and seal and assurance information treatment groups, with respondents having higher concerns in the seal and assurance information group (mean increment of .441; p = .04). In addition, we find a significant difference in the disposition to trust between the website and seal only treatment groups, with respondents having a higher disposition to trust in the seal only group (mean increment of .433; p = .03). We believe that these differences in our controls do not distort our findings. First, although concerns are higher in the seal and assurance information treatment, perceived assurance and trust perceptions are still higher compared to the seal and website treatment group. Therefore, one might assume that displaying assurance information is effective even in cases where users have high concerns. Second, although the disposition to trust is higher in the seal only treatment, the trust perception toward the system owner is not significantly different, which emphasizes the ineffectiveness of showing only the seal.

Table A.3. Mean Value and Standard Deviation of Constructs for Each Treatment Group

Group	**Intended effects mean (STD)**		**Control variables mean (STD)**			
	Perceived assurance	Trust perceptions	Cloud knowledge	Cloud involvement	Disposition to trust	Privacy concerns
Website (control group)	4.700 (1.31)	5.246 (1.08)	3.979 (1.56)	4.670 (1.74)	4.278 (1.55)	4.972 (1.36)
Website with seal	4.831 (1.41)	5.346 (1.17)	4.021 (1.66)	4.443 (1.90)	4.711 (1.42)	4.711 (1.46)
Website with seal and assurance information	5.601 (.98)	5.663 (.91)	3.719 (1.62)	4.299 (1.69)	4.559 (1.29)	5.152 (1.16)

A5. Forced Seal Analysis

In line with our previous findings, we did not identify a significant difference between the website and forced seal treatments as well as between the seal and forced seal treatment (**Table A.4**). Treatment comparisons show that a seal and assurance information outperform even the forced seal in regard to users' perceived assurance (M_{Diff} = .840; p < .01) and trust perceptions (M_{Diff} = .373; p < .01), thereby supporting our previous findings.

Table A.4. Results of Forced Seal Treatment

GROUP	**INTENDED EFFECTS MEAN (STD)**		**CONTROL VARIABLES MEAN (STD)**			
	PERCEIVED ASSURANCE	TRUST PERCEPTIONS	CLOUD KNOWLEDGE	CLOUD INVOLVEMENT	DISPOSITION TO TRUST	PRIVACY CONCERNS
WEBSITE WITH FORCED SEAL	4.760 (1.32)	5.290 (1.00)	4.448 (1.32)	4.629 (1.63)	4.366 (1.51)	4.925 (1.31)

Appendix B. Suggestion Phase: Design of Effective Certification Presentations

B1. Explaining Selected Kernel Theories

Cognitive Load Theory

Cognitive Load Theory (CLT) assumes that the human cognitive architecture is twofold. First, it consists of a working memory that is limited in capacity when dealing with novel information and, second, it consists of a long-term memory that virtually has an unlimited capacity (Sweller et al., 1998). Working memory is a set of mental resources that people use to encode, activate, store, and manipulate information while they perform cognitive tasks (Baddeley, 2003). The CLT emphasizes that working memory must inevitably be limited in capacity when dealing with novel and unorganized information because the number of elements that needs to be organized increases linearly but the number of possible combinations increases exponentially (Sweller et al., 1998; van Merriënboer & Sweller, 2005). That problem does not arise when dealing with information from long-term memory that is already organized in schemata (Sweller et al., 1998; Sweller, 2004; van Merriënboer & Sweller, 2005). Schemata describe organized patterns of thought or behavior that structure information and relationships among them (Chi et al., 1982; DiMaggio, 1997). Therefore, information is acquired at any time, then used for related problems several times, and finally stored in a long-term memory for retrieval in the future (Paas et al., 2004; Sweller et al., 1998). If the learning process has occurred over a long period of time, a schema might evolve by incorporating a large amount of information and various lower-level schemata (Sweller et al., 1998). After extensive practice schemata can become automated, thereby allowing learners to unconsciously process information with minimal working memory load, bypassing working memory capacity limitations (Paas et al., 2004; Sweller et al., 1998). Although working memory is shown to only process a limited number of elements at a time, the size, complexity, and sophistication of elements are not due to (complex) schema construction (Sweller et al., 1998). The working memory treats schemata as one element, thus, there are no apparent limits on the amount of information that can be processed (Paas et al., 2004; Sweller et al., 1998). Structuring information so that the learner can quickly develop and automate schemata to store in the long term memory enhances knowledge acquisition and performance, and has become the focus of CLT (Paas et al., 2004; van Merriënboer & Sweller, 2005).

Toulmin's Model of Argumentation

Toulmin (1958) proposes a model of argumentation in daily communication which consists of six interrelated argument elements that appear to be common and invariant across different field settings (Figure B.1). First, a *claim*: that is an assertion or conclusion put forward for general acceptance, having a potentially controversial nature. Second, *data*: that are statements specifying the particular facts, evidence or previously established beliefs about a situation to support the claim. Third, *backing*: acts as evidence such as statistics or expert opinions, and explains why a warrant should be accepted. Fourth, a *warrant*: that is a (hypothetical) general statement that justifies the reasonableness of moving from data to a claim. VerLinden (1998) argues that backing should not only be used for warrants, but used for backing data as well to explicitly express reasons to accept data as accurate. Fifth, a *qualifier*: that is an articulation of the degree of certainty (e.g., possibly, perhaps, probably) associated with a claim. Finally, a *rebuttal* that is an extraordinary or exceptional circumstance that can defeat the warranted claim (Toulmin, 1958). Figure B.1 depicts the elements, including backing for data and their interdependencies.

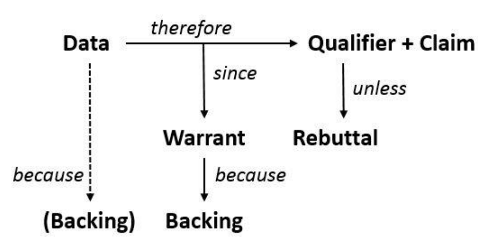
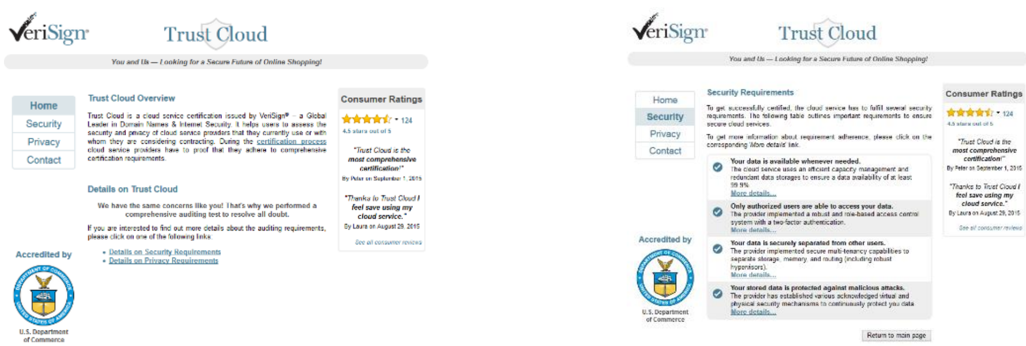


Figure B.1. illustration of Toulmin's model

B2. Survey Design to Validate the Applied Theoretical Model

We conducted a survey among 352 participants to validate the applied theoretical model. We built on the previous experiment and asked subjects to evaluate *MyCloudDrive* as an option for their personal use. A short description was given outlining the functions of *MyCloudDrive* and stating that the cloud service has been awarded the *Trust Cloud* certification. Subjects were told to inform themselves about the certification to reduce their concerns regarding privacy and security. Similar to the previous experiment, we then presented the subjects a website comprising assurance information related to the certification and asked the subjects to scrutinize the presented information. We implemented an interactive presentation that embeds different webpages to provide assurance information. On the first webpage, a short introduction to the certification was given, for example, stating that certified cloud services have proven to ensure security and privacy of a user's data. Subjects could also move along to different webpages that comprise information about security and privacy requirements that a system owner has to fulfil. Finally, we included some graphical cues (i.e., an emblem of the U.S. Department of Commerce) to increase the perceived realism of the given experiment scenario. Finally, we measured the constructs of interest. See Figure B.2 for screenshot snippets of the survey.



Interactive certification presentation with different webpages. For example, webpage (left) shows the main certification page and webpage (right) summarizes security requirements

Figure B.2. Snippets when validating the applied theoretical model (Suggestion Phase)

B3. Measurement

We reused the measures from our first experiment. We adopted the measures from D. Kim and Benbasat (2006) to measure argument quality. How to measure the multidimensional construct of the cognitive load has proven difficult for past researchers, resulting in multiple analytical approaches and related methods to measure the cognitive load (Paas et al., 2003; Sweller et al., 2019). In particular, rating scales have been widely used because people are quite capable of giving numerical indications of their experienced mental burden. In our study, we build on NASA's Task Loading Index (Hart & Staveland, 1988) and the related formative rating scale to measure users' cognitive load. Past research has shown these formative measures are reliable (cf. Gerjets et al., 2004; Paas et al., 2003; Speier & Morris, 2003). See Table B.1 for an overview of the items used.

Table B.1. Measurement Items Applied Theoretical Model

Cognitive load (Gerjets et al., 2004; Hart & Staveland, 1988)
* How hard did you have to work to accomplish the given task (to inform yourself about the Trust Cloud certification)?
* How much mental and perceptual activity was required (e.g., thinking, reading, deciding, remembering, etc.)?
* How much navigational activity was required (e.g., navigating through website, number of clicks etc.)?
[R] How successful do you think you were in accomplishing the goals of the task?
How much frustration did you have experienced during the task? (feeling insecure, discouraged, irritated, and annoyed versus secure, gratified, content, and complacent)
Argument quality (D. Kim & Benbasat, 2006)

How would you rate the certification information (e.g., regarding data protection and personal information usage) that you actually noticed in terms of the following adjectives?
* Formative unreliable items were dropped when testing the theoretical model. [R] reversed item

B4. Data Gathering

We recruited 352 participants from Amazon Mechanical Turk. We omitted 24 responses due to survey misuses (i.e., failed attention checks, did not read the assurance information, or rushed through the survey), resulting in a sample of 328 responses that fulfilled the initial data quality requirements.

B5. Data Analysis

The survey results were used to validate both the construct measurement scales and the proposed theoretical relationships. All indicators fulfilled the minimum loading requirements, demonstrating convergent validity. The calculations of CR demonstrated good internal consistency. Each construct showed an acceptable CR value greater than .7 (Nunnally, 1978). The AVE for each construct was greater than the suggested minimum of .5 (Fornell & Larcker, 1981). Since the cognitive load construct uses a formative construct specification, we checked for indicators' relative contributions and assessed calculated indicator weights and loadings as well as their significance. We dropped three cognitive load items because their relative and absolute contributions did not meet the required thresholds. Further, all variance inflation factor (VIF) values of the remaining cognitive load indicators were less than a threshold value of 3.3. Hence, the formative measurement model indicated low collinearity. We did not identify issues regarding discriminant validity, because the square root of each construct's AVE exceeded the inter-construct correlations (**Table B.2**), and the HTMT ratio of correlations showed values less than the value of .85 for all constructs.

Table B.2. Construct Reliability and Discriminant Validity

Construct	CR	AVE	Inter-construct correlations (HTMT)						
			(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Argument quality	.936	.787	.887 (-)						
(2) Cloud service involvement	.972	.919	.277 (.294)	.959 (-)					
(3) Cloud service knowledge	.945	.810	.107 (.111)	.586 (.623)	.900 (-)				
(4) Disposition to trust	.946	.815	.171 (.186)	.231 (.246)	.101 (.106)	.903 (-)			
(5) Trust perceptions	.966	.758	.605 (.647)	.190 (.197)	.001 (.035)	.256 (.266)	.871 (-)		
(6) Perceived assurance	.966	.780	.516 (.548)	.264 (.270)	.057 (0.57)	.265 (.279)	.757 (.786)	.883 (-)	
(7) Privacy concerns	.951	.707	.253 (.275)	.140 (.157)	.100 (.111)	-.123 (.128)	.059 (.070)	-.054 (.071)	.841 (-)
Note: cognitive load is omitted in this table due to is formative measurement.									

We account for CMB ex ante through the careful design of the questionnaire, applying the recommendations of Podsakoff et al. (2003), similar to our experiment in the problem awareness phase. Ex post, we run a full collinearity test and calculated the inner VIF value to control for CMB. The occurrence of a VIF greater than 3.3 is proposed as an indication of pathological collinearity, and also as an indication that a model may be contaminated by CMB (Kock, 2015; Kock & Lynn, 2012). All VIF values resulting the full collinearity test are lower than 3.3 (i.e., highest VIF value is 2.87), confirming that CMB is not a major issue in our data (Kock, 2015; Kock & Lynn, 2012).

Appendix C. Development and Evaluation Phase: Validating Certification Design Guidelines

C1. Experiment Procedure and Treatments to Test Design Propositions

We developed four instantiations of our design artifact (i.e., certifications presentations) to test the propositions in a 2x2 between-subjects experiment: (not) implementing DG1-DG3 and (not) implementing DG4. We randomly assigned subjects to one of the four treatments to detect the influence of each design manipulation (**Table C.1**). Although the certification presentations differed in their manipulations (i.e., applying respective design guidelines), all certification presentations had an identical style and layout. In this experiment, we reused the experimental procedure from the first experiment during the problem-awareness phase, and control for CMB (Podsakoff et al., 2003). The Figure C.1 for screenshot snippets for each treatment.

Table C.1. Overview of Developed Certification Presentation Instantiations

No.	Identifier	Manipulation	
		Cognitive load	**Argument quality**
1	**None (control group)**	high	low
2	**CP1**	low	low
3	**CP2**	high	high
4	**CP3**	low	high
Note: Gray-filled cells indicate manipulation			

Control Group (None): No design guidelines applied. The certification presentation shows a single block of text without paragraphs or text structuring. The text contains information about the certification authority, redundant information about the assurance requirements, security and privacy assurance claims, and information about the certification process. Thus, the certification presentation imposes a high cognitive load and has low argument quality.

Certification Presentation 1 (CP1): Reduce the cognitive load. To reduce the extraneous load and to manage the intrinsic load, we applied three design guidelines. First, to meet DG1, we reduced element interactivity by segmenting the certification presentation by content (i.e., we included headings, paragraphs, and lists to structure assurance information). With respect to DG2, we structured the certification presentation hierarchically and conveyed additional information in a hovering window (i.e., information about the certification process). Finally, we removed redundant and nonessential content to meet DG3.

Certification Presentation 2 (CP2): Enhance argument quality. We included in the certification presentation assurance arguments including claims, data, and backing to foster argument quality (in line with DG4). Overall, we included four security and four privacy arguments addressing users' potentially important cloud service concerns (Lansing et al., 2019; Trenz et al., 2018).

Certification Presentation 3 (CP3): Reduce the cognitive load and enhance argument quality. In this treatment, we manipulated *both* the cognitive load and argument quality by applying the two previous treatments simultaneously. To reduce extraneous overload and manage intrinsic load, only the argument's claim and data were shown. Subjects could view the argument's backing by clicking on a "*more details*" button next to each argument, resulting in a textbox fading in (fulfilling DG1 and DG2).

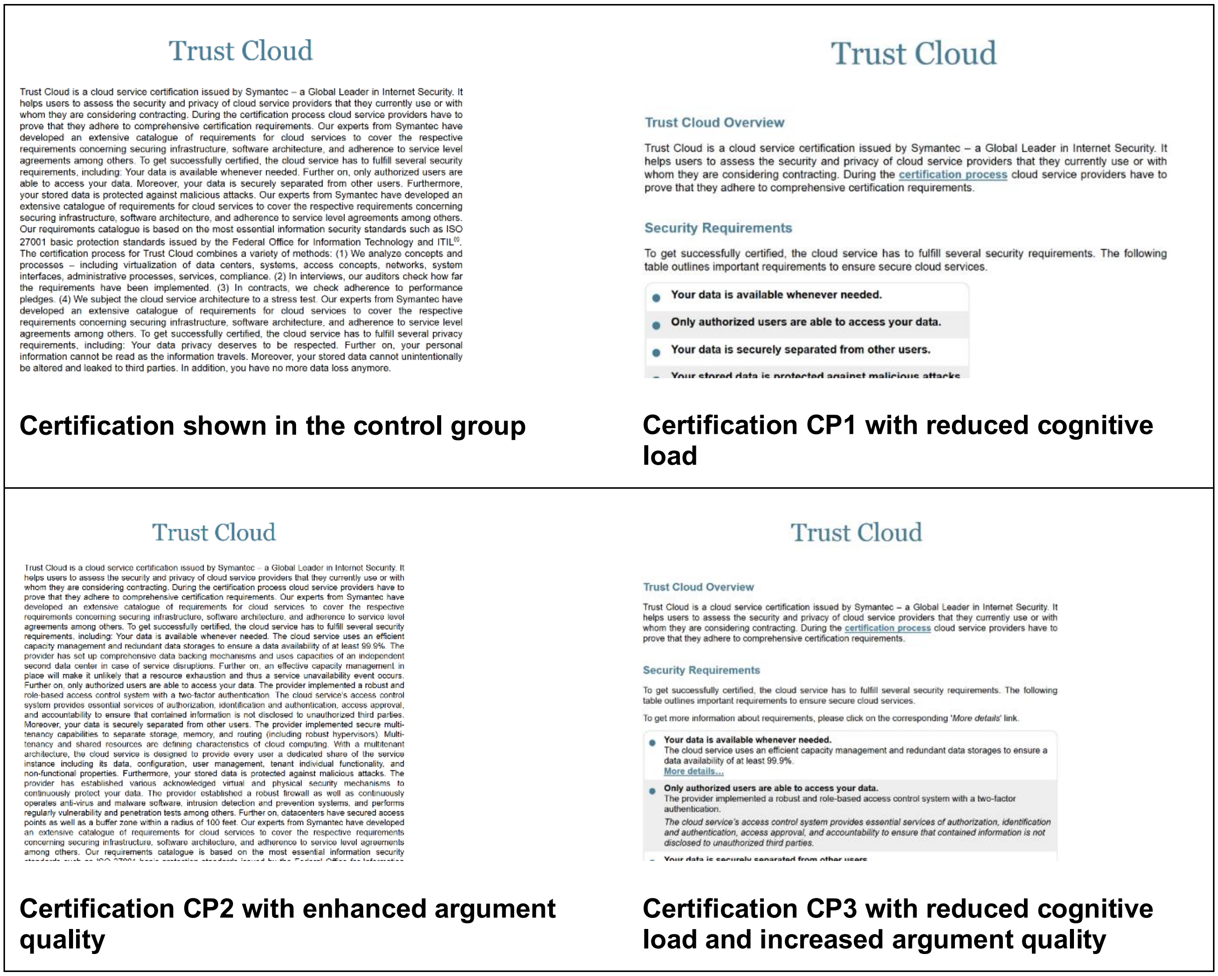


Figure C.1. Snippets when Validating Certification Design Guidelines (Development & Evaluation Phase)

C2. Pre-Test to Validate Treatment Designs

We tested whether we had successfully implemented design guidelines for certification presentation through a within-subject experiment. Here, 146 subjects from Amazon Mechanical Turk had to evaluate the four certification presentations in random order.

The pre-test experiment to test whether we successfully implemented our design guidelines consisted of five steps. After introducing subjects, we randomly presented a certification treatment and instructed subjects to carefully read the assurance information. We asked the subject a question concerning the certification's content to confirm a thorough reading (e.g., "*One privacy requirement highlights compromising of personal information. Therefore, the cloud service encrypts messages by the use of latest...(a) Advanced Encryption Standard (AES), (b) Secure Sockets Layer (SSL), (c) Temporal Key Integrity Protocol (TKIP), or (d) This information is not provided in the previous certification layout.*"). Fourth, we asked the subjects several questions to evaluate whether our manipulations were successful. The items were measured using seven-point Likert-type scales (strongly disagree-strongly agree). Manipulation check questions were carefully selected, discussed and pretested with 46 subjects, leading to minor adjustments (**Table C.2**). We randomly presented one of the remaining certification presentations to reduce order effects and repeated steps three through five until each certification presentation was presented and evaluated by the subject. The subjects were able change their responses for each certification presentation after they had seen another certification presentation to foster comparisons.

Table C.2. Manipulation Check Items

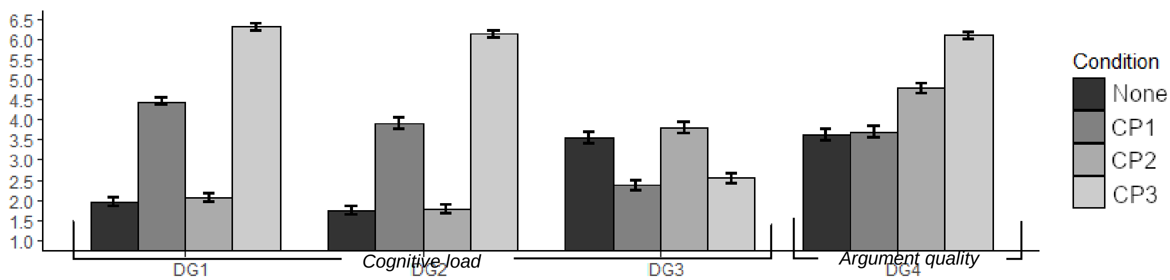
DG	Item
DG1	The certification information was subdivided into comprehensible parts.
	Each certification requirement (and related argumentation) was presented separately.
	The certification contained a number of links to reveal further information.
	Additional information was accessible via links.
DG2	The certification provided a clear and easy structure to navigate to (additional) information.
	Additional information was provided by links that do not impair reading fluency.
	Links were labeled appropriately to ease navigation.
	Additional information was shown in text boxes that optionally fade in by clicking on links.
DG3	In my opinion, the text contains redundant information.
	In my opinion, there are many duplicate statements in the text.
DG4	Detailed information was given for each individual security and privacy requirement.
	Each individual security and privacy requirement was backed up by additional data and arguments.
	The objective of each individual security and privacy requirement was reasoned in detail.
	Detailed information about what a cloud service provider has to do to fulfill each individual security and privacy requirement was given.

We ran an independent two-sample Welch's t-test to compare the certification presentations (**Table C.3**). Findings reveal that we successfully manipulated the certification presentation according to the design guidelines. The results of the pretest show that the certification presentation has the effects expected by the design guidelines (Figure C.2).

Table C.3. Manipulation Check Analyses

Comparisons	**Design guideline manipulation check variables mean difference (p-value)**			
	DG1	**DG2**	**DG3**	**DG4**
None *vs.* CP1	2.519 (< .01)	2.156 (< .01)	-1.178 (< .01)	.053 (.56)
None *vs.* CP2	.108 (.53)	.026 (.94)	.264 (.72)	1.158 (< .01)
None *vs.* CP3	4.378 (< .01)	4.389 (< .01)	-1 (< .01)	2.469 (< .01)

CP1 *vs.* CP2	-2.411 (< .01)	-2.13 (< .01)	1.442 (< .01)	1.104 (< .01)
CP1 *vs.* CP3	1.86 (< .01)	2.233 (< .01)	.178 (.09)	2.416 (< .01)
CP2 *vs.*CP3	4.271 (< .01)	4.363 (< .01)	-1.264 (< .01)	1.312 (< .01)
Note: Gray-filled cells indicate a significant difference between groups at p < .05				



Note that higher values of DG3 imply subjects perceived a significantly higher amount of redundant and nonessential content.

Figure C.2. Bar Graphs Showing Means and Standard Errors for Treatment Comparisons

C3. Data Gathering and Measurement

We next recruited 400 subjects from Amazon Mechanical Turk following the same restrictions as earlier experiments. We removed 28 responses from the final samples because the subjects failed the attention checks, did not read the certification presentation, or completed the experiment incorrectly (seven subjects in the control group, six subjects in the CP1 group, eight in the CP2 group, and seven in the CP3 group). This process resulted in 372 usable responses that fulfilled our data quality requirements. The subjects participating in the experiment were evenly distributed by gender (female 48%); were, on average, 37 years old (minimum 18 years, maximum 68 years); and had a high school (12%), college (25%), undergraduate (50%) or postgraduate degree (10%). Further, the subjects were employed (73%), looking for work (8%), students (6%), not looking for work (6%), or retired or unable to work (5%). The measurement instrument used the same scales as we applied while validating the theoretical model (Figure 3 and Table B.1).

C4. Statistical Analysis

We assessed the measurement scales for validity and reliability, and the scales showed good internal consistency, convergent validity, and discriminant validity (**Table C.4**). We run a full collinearity test and calculated the inner VIF value to control for CMB. All VIF values resulting the full collinearity test are lower than 3.3 (i.e., highest VIF value is 2.504), confirming that CMB is not a major issue in our data (Kock, 2015; Kock & Lynn, 2012). We ran a one-way MANOVA to test whether users' *cloud knowledge, cloud involvement, disposition to trust* or *privacy concerns* differed across treatments and found that all treatment groups were structurally equivalent for the control variables (F(12, 965.991) = 1.091, p = .364).

Table C.4. Construct Reliability and Correlations

Construct	CR	AVE	Inter-construct correlations (HTMT)				
			(1)	(2)	(3)	(4)	(5)
(1) Argument quality	.933	.776	.881 (-)				
(2) Cloud service involvement	.978	.935	.244 (.252)	.967 (-)			
(3) Cloud service knowledge	.945	.810	.107 (.104)	.599 (.636)	.900 (-)		
(4) Disposition to trust	.936	.787	.103 (.112)	.183 (.192)	-.016 (0.38)	.887 (-)	
(5) Privacy concerns	.941	.669	.372 (.392)	.162 (.171)	.150 (.163)	-.073 (.083)	.818 (-)
Note: cognitive load is omitted in this table due to is formative measurement							

We ran a MANOVA and diverse post-hoc tests to test whether the treatments significantly differed regarding cognitive load and argument quality. **Table C.5** summarizes the LDS post-hoc test results.

Table C.5. LDS post-hoc test results

Differences between	**Differences of Means**	**SE of Difference**	**95% CI**		**p-value**
			Lower	**Upper**	
Cognitive load					
None – CP1*	.301	.155	-.004	.606	.053
None – CP2 ***	-.779	.156	-1.086	-.473	.001
None – CP3**	.312	.156	.006	.618	.046
CP1 – CP2 ***	-1.080	.156	-1.386	-.774	.001
CP1 – CP3	.011	.155	-.294	.316	.944
CP2 – CP3 ***	1.091	.156	.784	1.398	.001
Argument quality					
None – CP1	-.065	.199	-.456	.326	.744
None – CP2	.213	.200	-.180	.607	.287
None – CP3 ***	-.562	.200	-.954	-.170	.005
CP1 – CP2	.278	.200	-.114	.671	.164
CP1 – CP3 **	.497	.199	-.888	-.105	.013

CP2 – CP3 ***	.775	.200	-1.169	-.382	.001
Germane load					
None – CP1 **	.56	.245	.08	1.04	.023
None – CP2 ***	-1.21	.246	-1.70	-.73	.001
None – CP3 **	.62	.246	.14	1.11	.012
CP1 – CP2 ***	-1.77	.246	-2.26	-1.29	.001
CP1 – CP3	.06	.245	-.42	.55	.795
CP2 – CP3 ***	1.84	.246	1.35	2.32	.001
Intrinsic load					
None – CP1 ***	.90	.235	.44	1.36	.001
None – CP2 ***	-.88	.236	-1.34	-.41	.001
None – CP3 ***	.70	.235	.24	1.16	.003
CP1 – CP2 ***	-1.77*	.235	-2.24	-1.31	.001
CP1 – CP3	-.20	.235	-.66	.26	.395
CP2 – CP3 ***	1.57	.236	1.11	2.04	.001
Extraneous load					
None – CP1 **	-.62	.215	-1.05	-.20	.004
None – CP2	-.29	.216	-.71	.14	.186
None – CP3 ***	-.58	.215	-1.00	-.16	.007
CP1 – CP2	.34	.215	-.09	.76	.118
CP1 – CP3	.04	.215	-.38	.47	.840
CP2 – CP3	-.29	.216	-.72	.13	.174
Note: * = p < .1; ** = p < .05; *** = p < .01					

Figure 1. Exemplary Presentations of IS Certifications as Assurance Seals

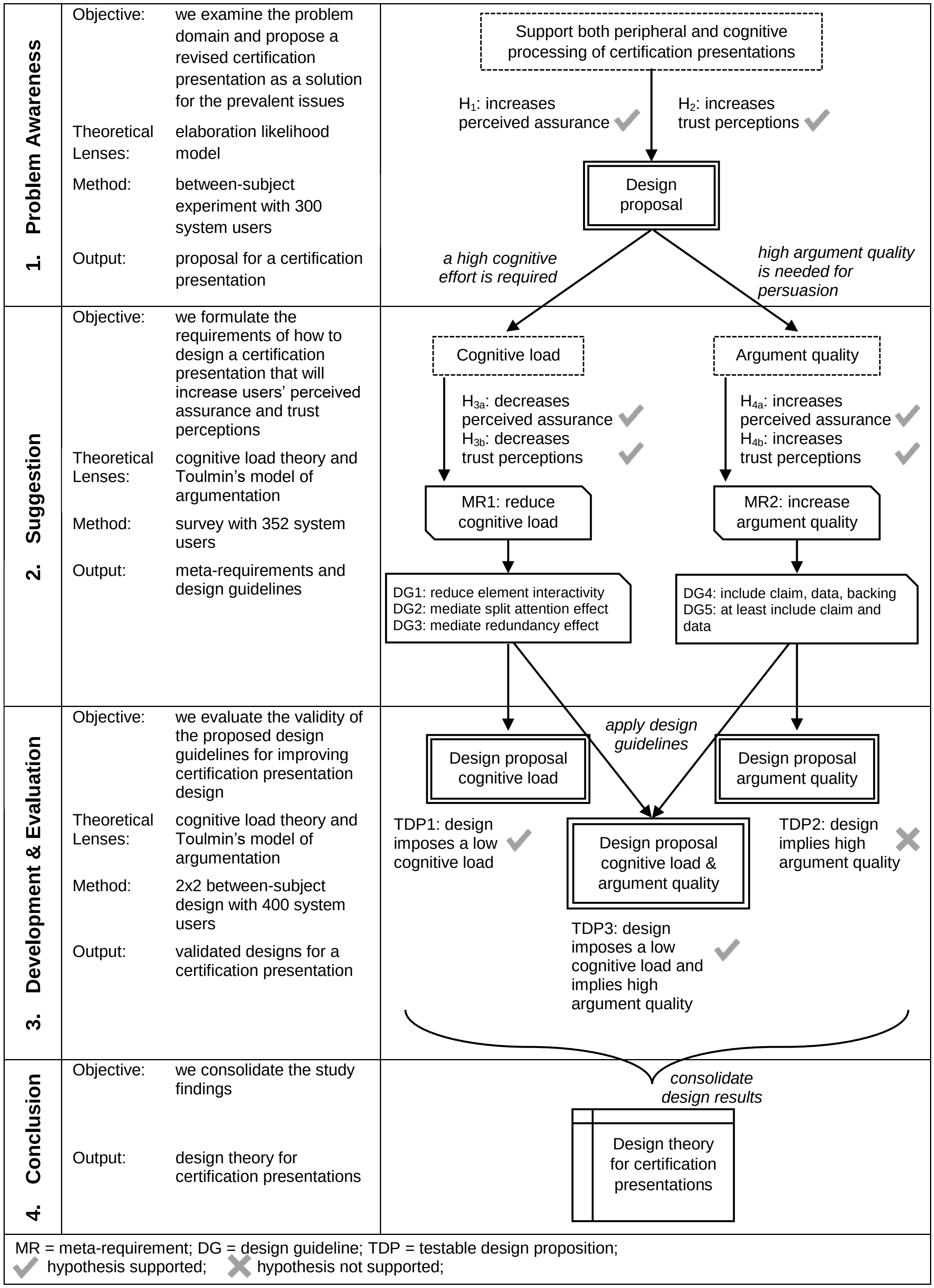


Figure 2. Summary of the Study Design and Research Phases

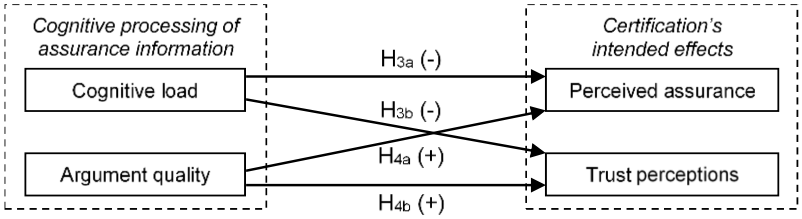
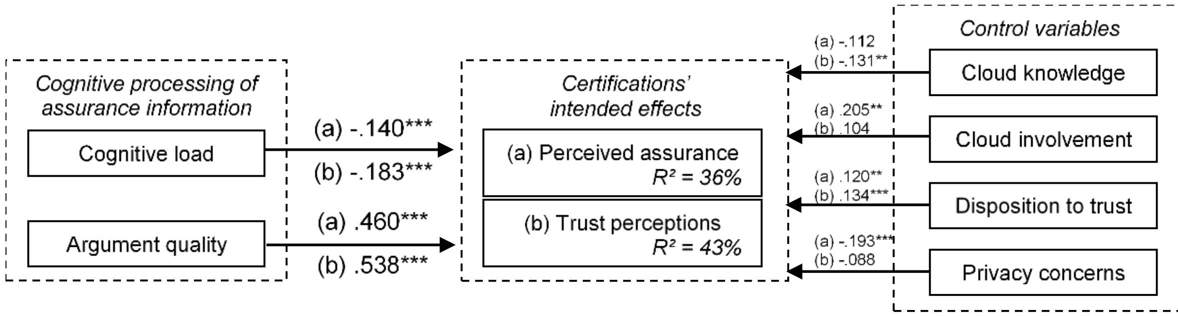


Figure 3. Applied Theoretical Model to Derive Meta-Requirements



Note: (a) perceived assurance, (b) trust perceptions; * = p < .1; ** = p < .05; *** = p < .01

Figure 4. Results of Theoretical Model Testing Showing the Direct Effects

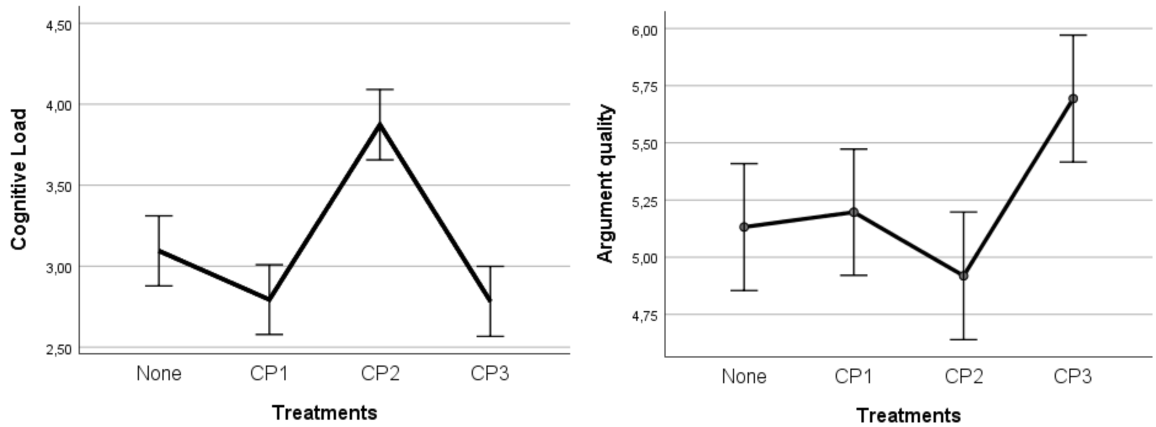


Figure 5. Graphs Showing Means and Error Bars for Treatment Comparisons

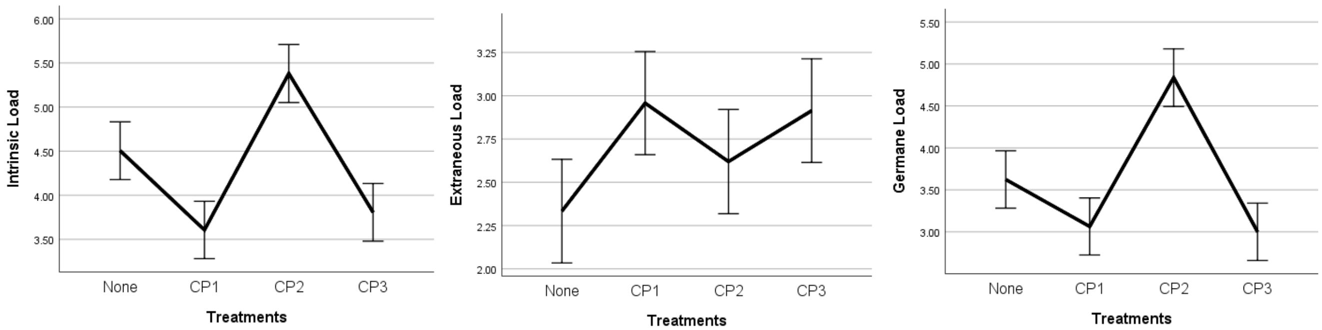


Figure 6. Graphs Showing Means and Error Bars for Treatment Comparisons on the Cognitive Load Types

Table 1. Example Research Analyzing the Effectiveness of Certifications

Study	Outcome measure		Context	Sample	Product category
	Perceived assurance	*Trust perception*			
McKnight et al., 2004		o	Service information	Students	Legal advice
Rifon et al., 2005	o	+	Online shopping	Students	Music
Yang et al., 2006	+		Online shopping	Students	Web cameras
D. J. Kim, 2008		M	Online shopping	Students	Various
D. J. Kim, Ferrin, & Rao, 2008		o	Online shopping	Students	*Unspecified*
Fisher & Chu, 2009		o	Online shopping	Students	Textbooks
Hu et al., 2010		M	Online shopping	Students	Various
K. Kim & Kim, 2011	+	+	Online shopping	Students	Running shorts
Lowry et al., 2012	o		Online booking	Students	Travel
Shah et al., 2014	o		Online shopping	Internet users	*Unspecified*
Hoffmann et al., 2015		+	Service selection	Internet users	*Unspecified*
Mavlanova et al., 2016		+	Online shopping	Students	Medicine
Löbbers & Benlian, 2019	M		Online shopping	Internet users	Computers
Siegfried et al., 2020		M	Online shopping	Internet users	Computers
Adam et al., 2020		M	Service selection	Internet users	Cloud services
Note:+ = positive, significant effect; o = no (significant) effect; M = under some conditions, the certification has an effect					

Table 2. Overview of the Study Design, Following Arazy et al. (2010)

This study phases	Problem awareness	Suggestion	Development and evaluation	Conclusion
Components of Arazy et al. (2010)'s theory-driven design framework	(Design problem), kernel theories	Applied theoretical model, meta-requirements, design guidelines	Testable design propositions	(Design theory)

Table 3. Study’s Major Terms and Their Definitions

Term	Definition
Certifications	Certifications are neutral third-party attestations of specific system characteristics, operations, and management principles to prove compliance with regulatory or industry requirements (Lansing et al., 2018).
Certification presentation (design artifact)	A scheme that provides the means to communicate certification-relevant information.
Perceived assurance	Perceived assurance refers to a user’s perception of the likelihood the system owner will try to protect the user’s confidential information collected during transactions and has applied security measures (D. J. Kim, Ferrin, & Rao, 2008).
Trust perceptions	Trust perceptions are defined as users’ perceptions of the competence, benevolence, and integrity of a system owner (McKnight et al., 1998).
Cognitive load	The cognitive load is referred to as the level of cognitive effort necessary to scrutinize and elaborate on the assurance information argument (Sweller, 1988).
Argument quality	Argument quality refers to a user’s perception that a message’s arguments are strong and cogent (Petty & Cacioppo, 1986).

Table 4. Key Issues Surrounding IS Certifications Identified in Related Research

Issues	Description	Example Studies
Unknown meaning	Users are incorrect or uncertain about the meaning of displayed IS certifications.	Fisher & Chu, 2009; Löbbers et al., 2020; Lowry et al., 2012; McKnight et al., 2004; Moores, 2005
Limited familiarity with assurances	Users have incomplete and inaccurate perceptions of a certification's assurances.	Kimery & McCord, 2006; Kovar et al., 2000; Löbbers et al., 2020; Lowry et al., 2012; Portz et al., 2000
Expectation gap	User have divergent expectations between the assurances the certification intends to provide and the system qualities users perceive.	Houston & Taylor, 1999; Lala et al., 2002; Lansing et al., 2019; Odom et al., 2002

Table 5. Meta-Requirements for the Certification Presentation Design and Associated Hypotheses of the Applied Theoretical Model

No.	Meta-requirement description	Associated model hypotheses
MR1	Reduce the cognitive load imposed by the assurance information embedded in certification presentations to enhance the level of perceived assurance and trust perceptions	H_3
MR2	Enhance the argument quality of the assurance information embedded in certification presentations to enhance the level of perceived assurance and trust perceptions	H_4

Table 6. Design Guidelines for the Certification Presentation Design and Corresponding Meta-Requirements

No.	Design guideline	Corresponding MR
DG1	Manage the intrinsic cognitive load by reducing element interactivity using segmentation	MR1
DG2	Mediate the split attention effect using a hierarchical structure, reduce the number of hyperlinks, and prevent the redirection of users to other pages	MR1
DG3	Mediate the redundancy effect by removing redundant and nonessential content, graphics, and subject actions	MR1
DG4	Assurance statements will have higher argument quality if they include the claim, data, and backing	MR2
DG5	Assurance statements will enhance argument quality if they include at least a claim and data	MR2

Table 7. Overview of Design Propositions

No.	Testable design proposition	Corresponding design guidelines
TDP1	It is feasible to design a certification presentation that imposes a lower cognitive load by employing segmentation, structuring elements hierarchically, using textual overlays for ancillary information and removing redundancies	DG1, DG2, DG3
TDP2	It is feasible to design a certification presentation that has higher argument quality by including in each assurance statement an argument's claim, data, and backing	DG4
TDP3	It is feasible to design a certification presentation that (1) imposes a lower cognitive load by employing segmentation, structuring elements hierarchically, using textual overlays for ancillary information and removing redundancies and (2) that has higher argument quality by including in each assurance statement an argument's claim, data, and backing	DG1, DG2, DG3, DG4

Table 8. Design Components for Developing a Design Theory for Certification Presentations

Component	Description
Core components	
1) Purpose and scope	System owners and certification authorities can enrich certification presentations with persuasive and informative assurance information to counteract current issues with seals only as presentations and to ensure they effectively convey complete and reliable assurance information. Certifications must be properly designed to create added value, provide as much information as needed by users, and increase users' perceived assurance and trust perceptions. Certification presentation designs should include enhancements for a low cognitive load and high argument quality.
2) Constructs	Cognitive load, argument quality, perceived assurance, and trust perceptions (Table 3)
3) Principle of form and function	We derive the following design guidelines for certification presentations: • Manage the intrinsic cognitive load by reducing element interactivity using segmentation • Mediate the split attention effect using a hierarchical structure, reduce the number of hyperlinks, and prevent the redirection of users to other pages • Mediate the redundancy effect by removing redundant and nonessential content, graphics, and subject actions • Assurance statements will have higher argument quality if they include the claim, data and backing • Assurance statements will enhance argument quality if they include at least a claim and data
4) Artifact mutability	Enriching mere graphical seals with assurance information is effective for systems exhibiting a high degree of user uncertainty and information need or that are operated by new or unknown system owners. Our research findings are applicable to both cloud and other online systems with similar characteristics and uncertainties. The generalizability and transferability of our findings may be limited because users may not have a pressing need to inform themselves about the system owner. Derived meta-requirements and design guidelines can be applicable in related web-assurance design domains, for example, when designing privacy policies or service-level agreements.
5) Testable propositions	We evaluate certification presentations designs that • impose a low cognitive load by employing a segmentation, structuring elements hierarchically using textual overlays for ancillary information and removing redundancies; • have higher argument quality by including in each assurance statement an argument's claim, data, and backing; and • impose a low cognitive load and have high argument quality.
6) Justificatory knowledge and kernel theories	We base our research on the elaboration likelihood model to understand each user's manner of information processing. To enhance cognitive processing, we build on cognitive load theory and Toulmin's model of argumentation.
Additional components	
7) Principles of implementation	Our findings reveal that certification presentations must enhance both the cognitive load and argument quality to achieve benefits related to a convincing argumentation structure.
8) Expository instantiation	We developed two certification presentation instantiations to test each design proposition; one instantiation to test the combined effects of a low cognitive load and high argument quality; and one instantiation incorporating no design guidelines (as a control group).

Table 9. Summary of Main Findings and Their Relation to Existing Research

Previous research gaps	This study's findings	Implications for research
Prior research analyzes the impact of certifications just from a black box perspective (i.e., embedded vs. not embedded a seal), and lacks a deep understanding about how users actually process certifications (Lansing et al., 2018).	Certification design proposal that facilitates users' peripheral processing by comprising a graphical seal and facilitates users' cognitive processing by providing detailed assurance information that richly informs users of underlying assurances.	We contribute to research by providing deeper insights into how users process and understand the nature and role of certifications and underlying assurances, leading to expected certification effects.
		This finding answers frequent calls of researchers to improve the design of certification presentations (cf. D. J. Kim, Steinfield, & Lai, 2008; Kimery & McCord, 2006; Lansing et al., 2018; Lowry et al., 2012; Lynn et al., 2016).
		Our research complements prior research (e.g., Lansing et al., 2018; Lansing et al., 2019) because it provides efficient means for providing persuasive assurance information to (potential) users.
Existing research has focused on analyzing the design of seals (e.g., Aiken & Boush, 2006; D. J. Kim, Steinfield, & Lai, 2008; Moores, 2005) but lacks guidelines on how to design certification presentations embedding assurances (Lynn et al., 2016).	Design knowledge about certification presentations in the form of a set of meta-requirements and design guidelines that overcome users' limited understanding of certifications and related assurance information.	We contribute to research by identifying cognitive load and argument quality as two major boundary conditions of certification effectiveness. With this our research helps to open prior research's black box perspective on users' processing of certification presentations and provides the foundation for deriving appropriate design guidelines.
		With the exaptation (Gregor & Hevner, 2013) of CLT principles, we reveal that effective certification presentations can control high element interactivity as well as split attention and redundancy effects while applying derived design guidelines.
		We extend prior research by deriving several components to formulate a design theory (Gregor & Jones, 2007) for impactful certification presentations and provide an explanation of why the design works.
		We inform research by performing a thorough evaluation of design guidelines for certification presentations that has been neglected.
While researchers have frequently applied Toulmin's model of argumentation in online contexts, prior research mostly modified a single argument (e.g., Racherla et al., 2012; Riedl et al., 2010; Yi et al., 2013) and has overlooked the downside of applying a complex argument structure.	New insights into the interdependencies between argument structuring and the cognitive load.	This finding offers new insights into the application of Toulmin's model of argumentation by revealing a boundary condition when having a need to provide several interrelated assurance arguments. Our findings emphasize that assurance arguments need to include a claim, data, and backing to enhance argument quality, but the cognitive load should be reduced simultaneously by organizing the statements per the design guidelines suggested.