Toward replication study types for design science research

Alfred Benedikt Brendel[1] iD **, Tim-Benjamin Lembcke[2]** iD **,
Jan Muntermann[2] and Lutz M Kolbe[2]**

Abstract

In design science research, two important challenges exist to achieve greater influence in research and practice: (1) foster frequent reuse of artifacts and design theories and (2) increase knowledge accumulation in the field. In this article, we argue that replication studies could support the accumulation and development of design theories to reach a state that encourages reuse of artifacts and design theories. However, it is unclear precisely how replication relates to design science research—that is, what outcomes replication produces and how researchers should apply it within design science research. This study proposes three overarching research questions (*Does the artifact provide utility? Is the design theory complete? What design theory components fit a larger context?*) and eight categories for replication studies in design science research (Test, Redesign, Justification, Adaptation, Explanation, Update, Recreation, and Meta-Replication). We offer guidance to researchers, editors, and reviewers on how to conduct replication studies in design science research and why such studies are so critical. Our goal is to provide "food for thought" on the significance of design science research replication studies and, in turn, help facilitate their widespread implementation and publication. We conclude our study by highlighting areas for further discussion and investigation, such as defining replication procedures and conceptualizing genuine replication goals within design science research.

Introduction

Scholars have recently criticized the design science research (DSR) community for producing artifacts and related design theories that are rarely reused (Kruse and Seidel, 2017; Vom Brocke et al., 2017). This lack of reuse has also fueled a discussion about knowledge accumulation in DSR. Vom Brocke et al. (2020) argued that many DSR studies currently stand on their own and do not build on existing design knowledge. This failure to reuse artifacts and design theories in different contexts limits the scope, extent, contribution, and effect of DSR studies. If theories are not reused, every contribution of design knowledge is effectively a standalone piece and fails to contribute to generalized theories.

In response to calls for replication in DSR (Niederman and March, 2014; Olbrich et al., 2017), we argue that the concept of replication may provide a helpful means of fostering reuse and knowledge accumulation within DSR. Through the replication of studies, confidence in a theory can increase or decrease (National Academies of Sciences

Engineering Medicine, 2019). A high level of confidence signals to researchers and practitioners that a design theory can provide a reliable blueprint for the real-world implementation of artifacts and a sturdy theoretical basis for developing new design theories. Hence, a high confidence level is a prerequisite for reusing and extending theories of any kind, allowing for knowledge accumulation (Vom Brocke et al., 2020). Therefore, replication may be an important building block supporting greater design theory reuse and knowledge accumulation in DSR.

Replication research also provides a means of filtering out unreliable theories (i.e. theories with a low level of

[1]Technische Universität Dresden, Germany
[2]University of Göttingen, Germany

Corresponding author:
Alfred Benedikt Brendel, Chair of Business Informatics, esp. Intelligent Systems and Services, Technische Universität Dresden, Helmholtzstraße 10, 01069 Dresden, Germany.
Email: alfred_benedikt.brendel@tu-dresden.de

associated confidence). The danger of not questioning existing study results and theories has been highlighted by the so-called "reproducibility crises" in other disciplines. A reproducibility crisis occurs when the results of many empirical studies are impossible or at least challenging to reproduce in subsequent studies. Most prominently, researchers in the field of psychology have reported that at least half of all psychological studies fail reproducibility tests, calling into question great portions of the discipline's knowledge base (Baker, 2015). In this context, Baker (2016) investigated reproducibility in various disciplines (chemistry, biology, physics, engineering, medicine, and environmental studies) by questioning more than 1500 researchers. Approximately 90% of the researchers perceived science to have at least a slight reproducibility crisis. Conducting replication studies in DSR would increase confidence in individual design theories and reduce the risk that ineffective design theories remain in the discipline's knowledge base.

In considering general replication discourse, related reviews (e.g. Gómez et al., 2010, 2014) have indicated that replication has focused on the replication of descriptive, explanatory, and/or predictive theories (Gregor, 2006) but not design theories. Considering work on replication in DSR, no prescription or guidance exists concerning replication research, albeit some have made general calls for replication research in DSR (e.g. Niederman and March, 2014; Olbrich et al., 2017). This study introduces the concept of replication in DSR in two ways. First, it offers reflections on the role that replication may play within DSR and the importance of replication for DSR. This section includes a discussion of how replication may address the reuse and knowledge accumulation problems of DSR by increasing or decreasing confidence in design theories. It also reflects on the fundamental differences when replicating theories that aim to provide "utility" rather than "truth." Second, this study develops eight replication study types that illustrate how different replication studies could be conducted in DSR. Specifically, it proposes a progressive structure between these study types to guide scholars through their DSR replication projects.

Research background

DSR

This section describes our epistemological positioning on DSR and illustrates the execution of DSR by following the combined framework of Hevner (2007) and Hevner et al. (2004). It concludes with a brief account of the nature of design knowledge.

Epistemological positioning. Defining the epistemological positioning of DSR can be quite challenging (Goldkuhl, 2012; Lee and Nickerson, 2010). This article supports the notion that DSR is pragmatic and focuses on utility over truth (Agerfalk, 2010; Goldkuhl, 2012; Iivari, 2015), thus investigating actions and their influence on environmental outcomes (i.e. utility). Pragmatism assumes that humans create meaning by interacting with a constantly changing and evolving environment. To achieve the desired environmental changes, actions can be motivated by purpose (e.g. relevance) and knowledge (e.g. rigor) (Goldkuhl, 2012). Pragmatism aims to facilitate an understanding of how actions can drive outcomes (e.g. change). This understanding provides utility by guiding decisions that support and enable actions to reach a desired outcome (Agerfalk, 2010; Goldkuhl, 2012; Hevner et al., 2004; Lee and Nickerson, 2010). As a result, research rooted in pragmatism provides prescriptive theories of design and action (Gregor, 2006; Gregor and Jones, 2007). This study follows Iivari (2015) and summarizes DSR with the following interrelated points:

1. DSR produces innovative and novel artifacts and design theories as its research outcomes.
2. The construction and evaluation of artifacts constitute the main research activities of DSR.
3. From an epistemological perspective, DSR seeks utility over truth, rendering it pragmatic in nature.

Research process. While DSR processes can be structured in many ways (Iivari, 2015; Leukel et al., 2014), the seminal framework of Hevner et al. (2004) remains the most frequently referenced structure (Brendel et al., 2018; Leukel et al., 2014). Consequently, we base our understanding of DSR implementation on a framework (Figure 1) that combines the frameworks of two seminal publications by Hevner (2007) and Hevner et al. (2004). This combined framework defines the general nature of DSR and facilitates the unification of the related descriptive vocabulary.

The DSR framework comprises three research cycles: relevance, rigor, and design. The relevance cycle connects design activities with their related sphere of action and practical environment, stimulating the convergence of design with the requirements of real-world problems. These requirements can originate from people, organizations, or technology. The relevance cycle introduces newly designed artifacts to the field. The rigor cycle connects design activities with existing research and knowledge bases, integrating and extending existing knowledge. The design cycle is at the center of the DSR model. This cycle represents the iteration of the construction and evaluation of artifacts. Artifacts can take different forms—constructs, models, methods, and instantiations—and their evaluation can be conducted using different methods, including observation, analysis, experimentation, testing, or description (Hevner, 2007; Hevner et al., 2004).

Regarding the overall contribution of a DSR artifact, scholars have identified four formats for extending information systems (IS) knowledge: (1) presenting new solutions for known
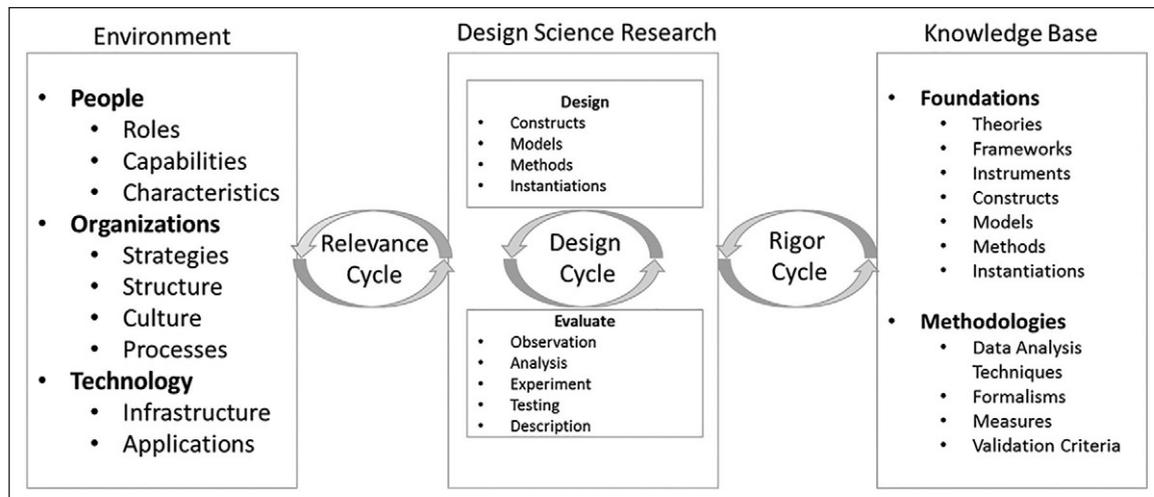
Figure 1. Design science research framework (based on Hevner, 2007; Hevner et al., 2004).

Table 1. Components of a design theory (Gregor and Jones, 2007).

Component	Description
I. Purpose and scope	"What the artifact is for": the set of meta-requirements or goals that specifies the type of artifact to which the theory applies and in conjunction also defines the scope or boundaries of the theory.
II. Constructs	Representations of the entities of interest in the theory.
III. Principles of form and function	The abstract blueprint or architecture that describes an IS artifact, a product, method, or intervention.
IV. Artifact mutability	The changes in the state of the artifact anticipated by the theory, that is, the degree of artifact change encompassed by the theory.
V. Testable propositions	Truth statements about the design theory.
VI. Justificatory knowledge	Underlying knowledge or theory from the natural or social or design sciences that provides a basis and explanation for the design (kernel theories).

IS: information systems.

problems, (2) extending known solutions to new problems, (3) proving known solutions for known problems, and (4) proving new solutions for new problems (Gregor and Hevner, 2013).

Design knowledge. Typically, a DSR project has two related but distinct outcomes (Baskerville et al., 2018; Beck et al., 2013): the artifact and an associated design theory. The artifact constitutes an instantiated solution for a relevant problem (Gregor and Hevner, 2013), whereas the design theory provides a summary of the gathered knowledge that relates to the problem characteristics and solution design (e.g. principles of form and function, requirements) (Gregor and Jones, 2007; Venable, 2006). The artifact provides the desired utility (Goldkuhl, 2012; Iivari, 2015), and the design theory offers an abstraction of the design, enabling the transfer of design principles (DPs), elements, or patterns (Gregor, 2006; Gregor and Jones, 2007; Gregory and

Muntermann, 2014) to different problem classes and instances (Iivari, 2015; Lee et al., 2011). Within this context, Gregor and Jones (2007) developed the anatomy of a design theory, which consists of six essential components (Table 1) and formalizes the design knowledge gathered during artifact development.

DSR contributes to research on how to solve certain problems by adding to the prescriptive and descriptive knowledge bases (Gregor and Hevner, 2013). Carrying out DSR adds solutions (e.g. components of the artifact) to the solution space, deepens the understanding of the problem space (e.g. requirements), and maps solutions to problems to provide utility (Figure 2).

This design knowledge can be found at different levels of abstraction (Gregor and Hevner, 2013; Lee et al., 2011). The first level is instantiation, which is defined as a situated implementation of an artifact and demonstrates the problem
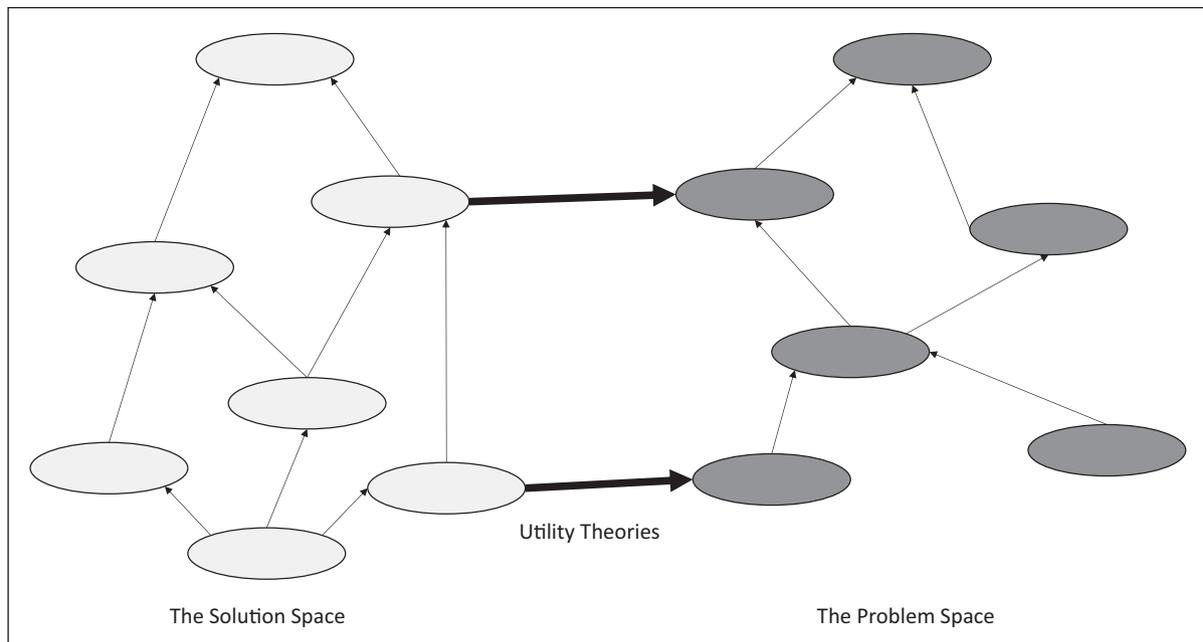
Figure 2. Solution and problem space relation (Venable, 2006).

solution in the form of a software product or implemented process. The second level of abstraction is nascent design theory (e.g. knowledge codified as operational principles or architectures). This level includes constructs, methods, models, and technological rules. At the highest level of abstraction, a well-developed design theory connects design knowledge with kernel and grand theories to facilitate knowledge of embedded phenomena (Gregor and Hevner, 2013). In this context, design theories provide a basis upon which to build and address new problems. However, as Vom Brocke et al. (2020) noted, existing design knowledge (i.e. design theories and DPs) is seldom reused, preventing such design knowledge from being abstracted, generalized, amplified, or contextualized.

Replication research

Within the greater context of scientific progress (e.g. scientific methods), replication plays the role of attempting to falsify existing theories. According to Popper (1959, 1963), theories must be falsifiable to be considered more than mere belief. Scientific progress can be conceptualized as comprising three acts: (1) proposing a theory, (2) refuting the theory, and (3) improving or replacing the theory to explain an investigated phenomena better (Salovaara and Merikivi, 2015). Going beyond Popper's vision of science, others have made the argument that confidence in theories increases if these theories remain unfalsified (i.e. attempts at replication are successful) (National Academies of Sciences Engineering Medicine, 2019).

Researchers conduct replication studies with at least one of five intentions: (1) finding sample errors, (2)

controlling for a lack of internal validity, (3) uncovering fraud, (4) expanding or generalizing results to cover a larger or different context, and (5) verifying original hypotheses (Schmidt, 2009). The aim is to investigate the overall generalizability and refine the results of previous studies, regardless of whether the findings are novel or surprising (Greulich and Brendel, 2019). Any theoretical contribution lies in the iterative improvement and elaboration of an existing theory, supporting it with empirical results or refuting it based on contradictory replication results (Compeau et al., 2012).

Overall, replication can be characterized as a "backward view," regarding what has been done, instead of what could be done (e.g. describing and explaining new phenomena) (Greulich and Brendel, 2019). Building on Dennis and Valacich (2014), who identified three replication study types (exact, methodological, and conceptual), Brendel et al. (2020) developed a set of six replication study types: (1) exact, (2) methodological, (3) context, (4) transfer, (5) method, and (6) comparison. They described and distinguished the study types along the dimensions of theory, method, and context (Table 2). In summary, all replication studies have an underlying structure that must share at least one aspect with the original study, be it the theory, method, or context.

Replication in DSR

In principle, each DSR artifact is fallible because it either does not work at all or does not work any longer (Lee and Hubona, 2009). In principle, this fallibility fulfills the requirement of being falsifiable (Popper, 1959, 1963) and

Table 2. Replication study categories (based on Dennis and Valacich, 2014; Brendel et al. (2020)).

Name	Theory	Method	Context	Description
Exact	Same	Same	Same	Replicating the same theory or results via the same method in the same context as the original study.
Methodological	Same	Same	Different	Replicating the same theory or results via the same method in a different context from the original study.
Context	Same	Different	Same	Replicating the same theory or results via a different method in the same context as the original study.
Transfer	Same	Different	Different	Replicating the same theory or results via a different method in a different context from the original study.
Method	Known	Same and Different	Known	Replicating a known theory or known results via different methods in a known context to validate the original method.
Comparison	Same and Different	Same	Same	Comparing the same theory or results with alternative theories or results via the same method in the same context as the original study.

Same: same as original study; Different: different from the original study; Known: known and different from the original study. "Known" means that the relation between theory and context has been well supported via various means.

supports the scientific character of DSR. Nonetheless, as in the general sciences, the opportunity to falsify or reject theories is not, in itself, enough, if other studies do not reassess and verify the theories (Popper, 1959, 1963; Salovaara and Merikivi, 2015). One way to conduct falsification research is by implementing replications. Replication studies can be a productive addition and can strengthen the scientific standing of DSR. The replicability of DSR and its artifacts and design theories raises questions about how such re-examinations may be conducted.

Deviations from traditional replication approaches are inevitable because the replication of truth-based theories is straightforward, particularly in the natural sciences (Dennis and Valacich, 2014) or computer sciences (Peng, 2011). In these areas, replications can rely on quantitative methods (e.g. validating p-values, the level of significance, or sampling errors) (Anderson and Maxwell, 2016; Bonett, 2012; William and Choong, 1999). In contrast, DSR prioritizes utility, which does not fit the existing replication types for three reasons.

First, DSR commonly follows an iterative research approach by combining various methods (Hevner, 2007; Hevner et al., 2004; Sein et al., 2011). In DSR, no method is defined in such a way that it could be replicated or applied to a different context. Even for the final evaluation of an artifact, subjective or contextual influences are likely to be present. The evaluation criteria may not be suitable in all instances of a problem that an artifact attempts to solve. For example, in specific contexts, the provision of timely results may be more important than a high level of accuracy, or vice versa.

Second, an extensive DSR study may include multiple DSR cycles, leading to attribution or causality problems. In this case, a researcher can no longer clearly attribute differing results in the replication study to specific methods or phases of the original study. This problem in replicability is due to the nature of DSR in addressing difficult problems, oftentimes failing to isolate causes and effects, as is more common in laboratory-like controlled research

settings. For example, in many circumstances, it is challenging to facilitate randomized controlled trials in DSR. Thus, replications have to be different in DSR because many influences may be present, and many may not be explicitly stated in the baseline DSR study. This problem reflects DSR's focus on utility over truth and its status as not being an exact science in the same sense as the natural or behavioral sciences.

Third, the environments and surroundings of artifacts are ever-changing and evolving (Gregory and Muntermann, 2014; Niederman and March, 2014), which—as a practical matter—renders exact replication difficult and perhaps even impossible. The evaluation of DSR is context-dependent and, frequently, does not address underlying cause-and-effect relationships but remains highly observational and time- and context-dependent regarding action-outcome relationships (Olbrich et al., 2017). For example, an artifact that provided utility 10 years ago may not provide any utility today (e.g. IS for DVD rental platforms) because of noteworthy shifts in user preferences and technological/competitive landscapes (e.g. transitioning to online streaming).

These factors render the framework of Brendel et al. (2020) and related replication study categories, including those of Dennis and Valacich (2014), inappropriate for DSR because its methodological approach and contextual circumstances differ. However, the fundamental concept of relating theory to reality via a research approach holds for DSR (Hevner et al., 2004; Iivari, 2015). An original design theory can be applied in a replication study, but the framework components of the method and context must be adapted to fit DSR, which leads to the need to formulate genuine DSR replications study types.

Research approach

At the beginning of this study, we screened and examined the present body of knowledge by using online databases to search for peer-reviewed articles on replication in DSR

(Appendix). That search led to the conclusion that replication in DSR has not been successfully implemented or conceptualized so far. In general, IS scholars have introduced and promoted the idea of replication research for DSR (Niederman and March, 2014; Olbrich et al., 2017) and have indicated that further replicative research is necessary.

For example, scholars have introduced the idea of "solution replication," which refers to replicating DSR by reexamining an artifact's utility in the context of an artifact's problem class (Olbrich et al., 2017). However, to the best of our knowledge, it appears that no completed DSR replication studies are available yet. There were also no seminal works on replications in DSR that go beyond stating the need for it and explaining its importance.

To facilitate the initial steps toward replication in DSR, this study suggests overarching research questions and replication types. These questions and types stem from the existing body of knowledge relating to replication research and DSR. Using a transdisciplinary bridge-building approach, this study transfers existing knowledge in other disciplines to the DSR context (Cooper, 1988). As replication must meet the specificities of DSR, we not only reviewed the literature but also transferred and adapted its concepts. Throughout the study, we readily sought feedback from IS and DSR scholars and discussed our collective understanding. For instance, the eighth replication type of meta-replication arose from scholarly feedback. In summary, this review and transfer revealed eight potential types of replication studies in DSR.

Derivation and proposition of eight replication study types for DSR

DSR replication research questions

As a starting point for replication in DSR, we suggest that replication in DSR, in general, aims to answer one or more of the following questions:

1. *Does the artifact provide utility?* Every DSR process aims to develop a novel solution for a relevant problem (Hevner et al., 2004). Therefore, the developed artifact should offer utility (Hevner et al., 2004; Iivari, 2015). However, DSR has largely omitted validation checks for falsification (Salovaara and Merikivi, 2015), and artifact evaluation is often based on the application within a single case (Arnott and Pervan, 2012; Leukel et al., 2014). Hence, one major DSR replication question seeks an improved understanding of the artifact's utility and its related design theory, aiming to confirm or disprove its utility for other problem classes.
2. *Is the design theory complete?* A typical design theory has six essential interconnected components

(Gregor and Jones, 2007). Each component consists of multiple elements (Table 1). For example, the "purpose and scope" component of a design theory includes a description of the design theory's goal (e.g. to solve a specific problem) and a list of related requirements. In this context, for example, the researchers performing the replication validate that the original list of requirements is complete, not lacking any important elements, and that the principles of design and function are fully formulated, covering all important aspects of design knowledge.

3. *What design theory components fit a larger context?* Researchers develop artifacts and design theories to provide or prescribe solutions for problem instances or classes (Iivari, 2015; Lee et al., 2011; Venable, 2006). By abstracting from gathered design theory insight, each artifact has implications for the IS discourse in large and related research streams, such as human–computer interaction or decision-support systems (Banker and Kauffman, 2004). To extend and provide validation of an artifact's design and utility, replication studies should address the application of the artifact in other environments to provide grounds for generalization. This process could reveal new elements of a design theory or suggest a re-evaluation of some of its elements. Connecting the artifact with kernel theories to explain its design also provides a valuable and worthwhile answer to this question (Gregor and Hevner, 2013; Hevner et al., 2004).

In summary, replication in DSR is not about developing a novel solution but about focusing on rigorously re-evaluating and extending existing solutions and theories.

Core considerations for replication in DSR

As outlined in section "Replication in DSR," the research process aspects of a DSR replication study and its relation to reality must be addressed. The conceptualization of the replication study types presented here rests on three main considerations:

1. *Design and evaluation.* The primary means of inquiry in DSR is to design and evaluate artifacts iteratively (Hevner, 2007; Hevner et al., 2004). DSR replication studies should not only point out the flaws of previous design theories but also propose possible adaptations or extensions. To facilitate scientific progress, the results of unsuccessful replications in DSR should include proposals for new design theories, extending or providing substitutes for the former proposals.

2. *Divide and conquer.* A DSR process can be characterized by an iterative and multi-method research approach (Gregory and Muntermann, 2014; Hevner, 2007; Hevner et al., 2004). Various methods can be combined to elicit requirements, gather knowledge, develop artifacts, evaluate solutions, and formulate design theories. As an attempt to replicate entire DSR projects may be impossible due to their complexity and various dependencies, implicit contexts, or the like, replication studies in DSR should initially focus on specific elements of previous DSR studies. For instance, solely addressing the elements of a study related to rigor, relevance, or design can be expected to lead to more concise and focused results than addressing all elements at once.

3. *Utility is a moving target.* Every test of the utility of an artifact is just a "snapshot" of the environment present at the moment of observation (Olbrich et al., 2017). Through rapid technological developments (Niederman and March, 2014), environments may change, leading to a lack of commensurability (Olbrich et al., 2017). Therefore, evidence in favor of or against the utility of a design theory must be distinguished from "noise" (Herwix and Rosenkranz, 2018) caused by differences in problem instances and changes in the environment. Hence, contradictory evidence should be analyzed via different means and from different perspectives. Building on the "divide and conquer" principle, contradictory evidence arising from a rigorous replication should be further investigated in the context of an artifact's relevance and/or design.

Development of study types

Following the first consideration (design and evaluation), we build upon the three DSR cycles presented by Hevner (2007) and Hevner et al. (2004) to formulate the research activities of a DSR replication. Each of the three cycles (rigor, relevance, and design) incorporates a different set of activities related to the DSR research process, which facilitates dividing an entire DSR project into manageable parts. Hence, based on the second consideration, the types were constructed by permuting the three DSR cycles (Hevner, 2007; Hevner et al., 2004) and using all eight potential combinations reveals an initial outline and set of potential replication types (Table 3). Furthermore, the cycles of relevance, design, and rigor suggest three approaches of replication:

1. Applying an artifact implemented according to the original design in practice, testing its utility (relevance).
2. Comparing the original artifact to a potentially improved version of it (design).

3. Relating the original design to the knowledge base, aiming to find contradictions (rigor).

Based on the design theory components of Gregor and Jones (2007), we identify and formalize the outcome of replication in DSR (Table 3). Each component can be addressed by one of the research activities, which may lead to changes to the original design theory. For example, a *test* replication study is concerned with activities related to the relevance cycle (e.g. understanding the environment, identifying business needs, gathering requirements, and applying the developed artifact). Consequently, a *test* replication provides insight into the purpose and scope (e.g. the requirements and goals) and the testable propositions (i.e. truth statements about the effects of an artifact built according to the design theory) of a design theory. Another example is a *justification* replication study, which concerns the rigor cycle (e.g. applying knowledge to or adding knowledge from prescriptive and descriptive knowledge bases). Such a study can provide implications for the design theory components of constructs (i.e. entities and agents related to the design theory) and justificatory knowledge (e.g. knowledge that provides explanations for the design). Eventually, changing or not changing the original design theory—and the related artifact design—interacts with the associated level of confidence. For instance, in the context of a *redesign* replication study, three different effects on the level of confidence can be distinguished:

1. The changes made to the original design provide no improvements and are considered unnecessary, increasing confidence in the original design.
2. The changes to the original design indicate improvements, leading to an updated version of the original design, in which someone can have higher confidence in the updated design compared to the original one.
3. Neither the improved nor the original design provides utility, reducing the level of confidence.

To address the third consideration (utility as a moving target), we implemented a progressive structural framework for replication study types. Because the context in which the research approach is applied is always somewhat different, researchers may encounter unexpected events. For instance, the intermediate results may indicate that the original design theory was not adequately formulated (e.g. results provide no utility or contradict established knowledge). To verify such results, the transition to a study type that addresses two or three cycles may be necessary and advisable. In summary, a DSR replication project can progress from a Level I type (addressing a single research cycle) to a Level II type (addressing two research cycles) or, eventually, to a Level III type (addressing all three research cycles), and this progress is guided by the results of each previous level (Figure 3). For example, a research

Table 3. Types of replication studies in design science research and their association with related research cycles and design theory components.

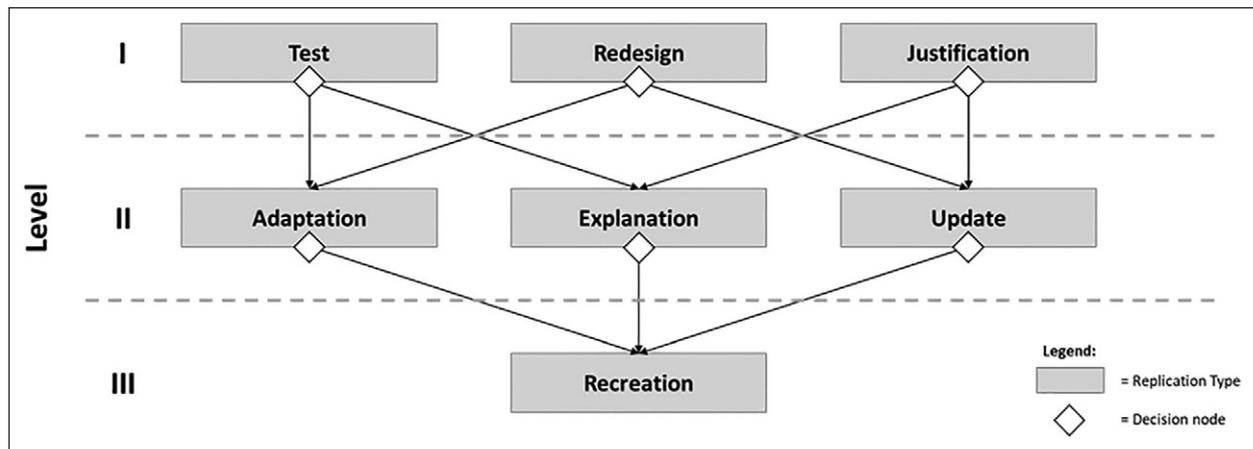| No.	Type	Replication activity			Replication outcome					
		Research cycle			Potential change in design theory components					
	Name *Key activity*	Relevance	Design	Rigor	I	II	III	IV	V	VI
					Purpose and scope	Constructs	Principles of form and function	Artifact mutability	Testable propositions	Justificatory knowledge
Level I										
1	Test *Prove utility*	X	–	–	X	–	–	–	X	–
2	Redesign *Ideate additional designs*	–	X	–	–	–	X	X	–	–
3	Justification *Generalize findings*	–	–	X	–	X	–	–	–	X
Level II										
4	Adaptation *Increase utility*	X	X	–	X	–	X	X	X	–
5	Explanation *Elucidate design*	X	–	X	X	X	–	–	X	X
6	Update *Suggest design*	–	X	X	–	X	X	X	–	X
Level III										
7	Recreation *Redo entire study*	X	X	X	X	X	X	X	X	X
Meta-level										
8	Meta-replication *Review artifacts*	–	–	–	X	X	X	X	X	X

Figure 3. Progression framework for non-meta DSR replication study types.
All non-meta-replication studies start as a test, redesign, or justification type.

process could begin with the intention of testing a design theory and its purpose, scope, and testable propositions (*test* replication study, Type 1). However, the replication may reveal that some testable propositions were not fulfilled, which calls for adjustments (e.g. changing the principles of form and function). Thus, the replication type progresses to an *adaptation* (Type 2) replication study.

In this context, the *meta-replication* study type is not part of the possible transitions because it does not contain a research activity related to the DSR cycles (e.g. relevance, rigor, and design). Conducting a meta-replication study is driven by the observation that practitioners are developing and implementing artifacts that have yet to be formalized as a design theory that could potentially inform future research.

To identify the appropriate Level I replication study type (Table 3 and Figure 3), scholars should consider their desired research outcome. Specifically, the expected answer to one of the three questions proposed in section "DSR replication research questions" (Does the artifact provide utility? Is the design theory complete? What design theory components fit a larger context?) should be matched with the study type. For example, if one expects an artifact—implemented according to a design theory—to not provide a sufficient level of utility, a *test* replication study should be conducted. Similarly, if one expects that a design theory lacks references for important justificatory knowledge, conducting a *justification* replication study would be appropriate. To aid in this assessment, we added an example subquestion to each type.

Illustration of replication types and progression framework

The following subsections describe the replication study types for DSR. Thoughts on how to implement each replication study type supplement these descriptions. For each replication type, a lead question is formulated that must be addressed and answered by a corresponding replication study. Answering these questions is decisive regarding the progression to another study type or the conclusion of the replication research process.

This section draws on a peer-reviewed DSR study (Seidel et al., 2018) to provide examples for each study type. In this example study, based on the concept of salient affordance, researchers developed DPs for IS that support organizational sensemaking. They performed three iterations of developing, demonstrating, and evaluating a prototypical implementation. The DPs were intended to prescribe how IS must be designed to support essential sensemaking practices in an organization's transformation toward environmental sustainability. The DPs were the following (Seidel et al., 2018: 245):

"DP 1: Provide novel information in the form of environmental facts, observations or general behaviour, so that the system affords users disruptive ambiguity and surprise in environmental sustainability transformations."

"DP 2: Provide functions of storing and simple and unambiguous categorisation of ideas, so that the system affords noticing and bracketing to users in environmental sustainability transformations."

"DP 3a: Provide features for interactive communication, so that the system affords users to engage in an open and inclusive discussion in environmental sustainability transformations."

"DP 3b: Provide users with an overview of all other users along with features for direct communication between users so that the system affords users to engage in an open and inclusive discussion in environmental sustainability transformations."

"DP 3c: Provide features to relate comments to other comments, so that the system affords users to comprehend circumstances and turning them into words and categories on a social ground in environmental sustainability transformations."

"DP 3d: Provide features to assign roles to users so that the system affords user-specific actions, such as moderation of discussions in environmental sustainability transformations."

"DP 4a: Provide features for categorisation of action possibilities to distinguish presumptions from actual planned actions, so that the system affords users presumption and action planning in environmental sustainability transformations."

"DP 4b: Provide features for dedicated feedback about the implementation and consequences of the implementation of actions in environmental sustainability transformations."

The researchers instantiated these DPs as an online collaboration and discussion platform. This platform allowed users to develop, communicate, and plan environmentally sustainable actions (e.g. replacing disposable plastic cups with reusable ones). The researchers conducted multiple evaluation rounds of the platform and corresponding DPs within a university with approximately 1000 students. The university provided an appropriate evaluation setting because the organization had recently begun a sustainability transformation that involved management, research teams, staff, and students. In the first round of evaluation, 51 users participated by using the platform over 2 weeks, including two focus group discussions at the end. The improved version of the artifact was evaluated in a second round by 99 users over 12 days and, again, included two focus group discussions. The last evaluation round consisted of a focus group discussion with four users. Seidel and his colleagues' study provides two main contributions to the field. First, it offers a set of theory-ingrained and empirically refined DPs for IS that support sensemaking in sustainability transformations. Second, it highlights how the concept of affordances could be applied within a DSR project to develop artifacts to support organizational practices.

The study of Seidel and his colleagues serves as an illustrative example of how different replication studies could be conducted in each of the eight study types. All of the following examples of how this study could be replicated and the potential results are entirely indicative and have no foundation in any conducted replication studies.

Type 1: test

Does the design theory provide utility? The main goal of DSR is to provide novel solutions for prevailing problems (Hevner et al., 2004; Winter, 2008). As such, solutions

should provide utility (i.e. solve the addressed problem) (Iivari, 2015). In a *test* replication study, scholars introduce an artifact into a similar application environment, attempting to reproduce the artifact's previously reported success in solving a class of problems. This can include to implement the original artifact, or an artifact built according to the original design theory. Researchers could discuss the artifact and its associated design theory with practitioners in, for example, interviews or workshops. The differences in problem environments (e.g. efficiency may be important in one case and usability in another) may lead to different perceptions regarding the usefulness of the artifact. These distinctions should not be attributed to "noise" (Herwix and Rosenkranz, 2018) but serve as an essential input to analyze the elements of a design theory, which may be meta-elements or may be specific to a particular context (Iivari, 2015; Lee et al., 2011). Regarding its outcome, a *test* replication study provides insight into the purpose and scope (i.e. a better understanding of the goal and requirements) and the testable proposition of a design theory (i.e. testing whether an artifact built according to the design theory fulfills the requirements).

If the original artifact design (e.g. a design theory) proves useful, the replication can be considered successful and can be concluded. If the application of the artifact design does not provide the expected level of utility (i.e. one or multiple testable propositions are not fulfilled), a transition to the *adaptation* type could be necessary to improve the design. Similarly, if the application of the artifact leads to unexpected results or unexpected behavior by users, a subsequent *explanation* replication may be needed.

In the example study, the sensemaking support system was tested in a university setting with approximately 100 participants. One replication approach could be to implement an artifact as described in the original study and to evaluate its utility in different organizational contexts, including companies and public organizations. This replication could be conducted via a field experiment to analyze the fulfillment of the contextual requirements and the perceived success in stimulating a sustainability transformation (e.g. evaluating the fulfillment of the testable propositions). For instance, during the replication study, employees did not engage in using the artifact due to time restrictions or the fear of openly exchanging potentially sensitive or controversial ideas (because environmental sustainability is a politically charged topic), which focus group interviews revealed. This result might be attributable to the setup of the original study as an experiment with students. A conclusion could be drawn that the initial design theory must be extended by adding functions to motivate employees and prevent the fear of stating controversial ideas (e.g. adding new requirements for the purpose and scope component). Addressing newly discovered requirements for adapting the present design could lead that study type to evolve into the *adaptation* type. If researchers

discover no need for adaptation, the progression to a Level II study type is not necessary.

Type 2: redesign

Is the design theory complete? Within the design cycle, two parts combine to develop an artifact as a solution. One part is the requirements and problem descriptions that stem from a relevance cycle. The other part is the corresponding and reviewed knowledge that stems from a rigor cycle. To replicate an artifact and its related design theory in a reconducted design cycle, scholars must consider the original inputs from the environment (relevance cycle) and knowledge base (rigor cycle). If the replicated design cycle leads to results similar to those of the original study, the replication succeeds and supports the original study results. Regarding the design theory, this result indicates that the replication has led to a similar set of principles of form and function (i.e. the artifacts are very similar) with a similar artifact mutability (i.e. the artifacts can be used in similar circumstances).

If the inputs from the original relevance and rigor cycle produce a substantially different artifact—compared with the original one—further investigation may be appropriate. Researchers should analyze the cause for the deviation, thus transitioning the replication study to *adaptation*. If the replication of the design cycle suggests that additional knowledge is necessary—that is, the researchers realize they are unable to leverage the original knowledge to determine a design that suffices—they should transition their replication efforts to an *update* study type.

In the example study, all authors were IS researchers. They shared an IS-related research bias in that they were all focused on knowledge from the fields in IS: computer science, management, and software engineering. When the original problem and its corresponding body of knowledge are handed off to a research team with completely different skill sets or worldviews, a substantially different artifact and design theory may arise. For instance, a team of social psychologists or a team of behavioral economists may bring a completely different sentiment to the proposed solution (e.g. focusing much more on influencing or stimulating measures). In the initial study, the final DPs appeared rather technical in terms of providing features for the categorization of actions or in assigning roles to users (DP 4a and DP 3d, respectively). A replication in the hands of a more psychologically focused research team might produce additional or different DPs. For example, the framing of the DPs could be less dependent on the detailed technological offering and more dependent on the circumstances surrounding the platform. DPs could be introduced that stimulate certain emotions or feelings, such as making users feel safe and comfortable when communicating on a platform. This goal could be realized by providing a trustworthy experience and by informing users whether, how, and by

whom their data could be viewed. In the case where researchers propose many new principles or changes to the original artifact (e.g. more than just minor adjustments, which could be the content of a thorough discussion), an evolution into an *adaptation* is necessary to evaluate the ideated solution design fully. Similarly, if an extensive consultation of other knowledge bases is necessary (e.g. theories from social psychology), the study type could evolve into an *update* type.

Type 3: justification

Is the design theory well justified? To reach the highest level of abstraction and maturity (Level 3: a well-developed design theory), a design theory must be connected to kernel theories and explain why the artifact is effective (Gregor and Hevner, 2013). However, finding and applying a high-level kernel theory to explain and justify an artifact's design is a challenging task. Therefore, some DSR researchers have opted to develop sophisticated instantiations instead of formulating design theories. To address this theoretical limitation, this replication study type focuses on the justification of a design, leading either to additional support for the design theory through furthering its theoretical foundations (i.e. validating the design theory components of justificatory knowledge and constructs) or to the disclosure of inconsistencies between the design theory and related kernel theories. If the replication study supports and theoretically justifies the original study, the replication endeavor concludes. Otherwise, a transition to an *explanation* type of replication study may be necessary, entailing a field validation due to the lack of evidence for the added justificatory knowledge. Likewise, a transition to an *update* replication study type might be beneficial if the knowledge base suggests an adaptation of the original study, meaning that an existing theory provides no support or limited support for the original design.

The example study researched organizational sensemaking. With its initial focus, the evaluation of the study was limited to small organizations and non-corporate environments. When aiming to generalize a theory, on a practical level, replication endeavors may concentrate on dissolving these limitations. In parallel, the theoretical level may be further substantiated by increasing the predictive power of the study. On one hand, current DPs could be synthesized into a more mature theoretical framework by providing a more abstract perspective on environmental and organizational sensemaking. Naturally, some trade-offs can occur, such as disregarding certain details that are too particular to specific implementations or problem instances. On the other hand, other meta-theories may be incorporated into a *justification* replication, leading to a more generalized theoretical contribution. If the researchers address the phases antecedent to individual decision-making, they may presume disruptive ambiguity and surprises to be relevant affordances to stimulate individual and collective behavior eventually. Replacing such instantiated affordances with a more general

concept, such as the self-determination theory (Deci et al., 1991; Ryan and Deci, 2004), a replication study may reevaluate the DPs and their effects concerning extrinsic motivation, intrinsic motivation, or amotivation.

One result could be that the current DPs stimulate individuals not previously intrinsically motivated by the goal (i.e. fostering environmental sustainability transformations). For people already willing to participate in these transformations, the current DPs may even translate into detrimental results (e.g. if the participants receive rash or destructive feedback for their suggestions on the platform). As a result, a broader design theory of motivational-inclusive sensemaking support systems might be derived. If the original design must be changed, the replication study might evolve into an *update* type of replication study. Furthermore, if the original DPs are expected to work at a higher level of abstraction, an evolution to an *explanation* replication study might be necessary to evaluate this assumption fully.

Type 4: adaptation

Does the design theory provide utility, and is it complete? Originating from an unsuccessful *test* or *redesign* replication, the adaptation type combines the DSR cycles of relevance and design. A researcher could decide to leverage the gathered insight to improve the original design, thus reestablishing the utility of an artifact. Ideally, scholars adapt the artifact until it can entirely solve the problem (again). During this process, the fit of a design theory can be evaluated, and additional requirements may be discovered. The researcher can then address these requirements, in turn, either by improving the generalizability of the original design theory or by developing a new version of the design theory.

Moreover, researchers can adapt the requirements to the new context's peculiarities, for instance, by reformulating or adding DPs. If successful, the replication endeavor ends here. If the original knowledge base cannot provide the necessary foundation to improve the design sufficiently, a transition to a *recreation* replication study type is advisable. As a combination of the *test* and *redesign* types (i.e. including relevance and design cycles), this replication study type addresses the design theory components of purpose and scope and testable proposition. The processes of adaptation and extension address the principles of form and function and artifact mutability components.

When considering our example study, the DPs for sensemaking support systems that facilitate sustainability transformation could be applied to non-sustainability contexts (e.g. stimulating corporate innovation management). For instance, DP 3b might be inappropriate for reaching the desired outcome because users are more personally involved in the related outcome (e.g. changes in sustainability measures may not lead to layoffs—organizational transformations, however, could do so). Hence, the artifact must be adapted, for instance by including a function for users to interact via pseudonyms to prevent negative repercussions for individuals.

Type 5: explanation

Does the design theory provide utility, and is it well justified? After an unsuccessful *test* or *justification* replication, the *explanation* type combines the rigor and relevance cycles to explain new field observations and insights further—or to provide evidence for the explanations or justifications. This type of replication study focuses not only on implementing the artifact but also on connecting it with a (new) knowledge base to explain the observed behavior of the artifact and its users in greater detail. As a result, an *explanation* replication can demonstrate that the knowledge base and field observations match; thus, the replication process concludes. In the case that the design itself does not seem to be effective or efficient any longer and must be adapted, a transition to a *recreation* study is appropriate. As a combination of the *test* and *justification* types (i.e. including the relevance and rigor cycles), this type of replication study primarily addresses the design theory components of constructs, justificatory knowledge, purpose and scope, and testable propositions.

In the example study, the theoretical background is focused on organizational sensemaking. The authors demonstrate how the sensemaking kernel theories of this study can be used to derive DPs for sensemaking support systems. In expanding the breadth of the proposed DPs and theory, frameworks such as the belief-action-outcome of Melville (2010) may provide additional theoretical perspectives. In practice, an *explanation* replication study may find evidence that some users increasingly participate over time while other users lose interest. This finding may challenge some of the proposed DPs (DP 2 to DP 4), which rely on users remaining interested, active, and engaged on the web platform. If scholars can demonstrate that the formation of heavy usage patterns coincides with users' beliefs about sustainability (i.e. those that have formed clear beliefs about corporate sustainability are more prone to active, ongoing engagement), an additional DP may be necessary to target the belief formation of employees with weaker sustainability beliefs. With the addition of these DPs, the study type evolves into *recreation*.

Type 6: update

Is the design theory well justified and complete? Stemming from an inconclusive *redesign* or *justification* replication, the *update* type combines the design and rigor cycles of DSR. With new insights present in a knowledge base, researchers may suspect that an existing artifact or design theory no longer works. For instance, a *justification* replication study could indicate the need to validate the proposed explanations of the original design, or a *redesign* replication study could require further grounding in a knowledge base. If an extended knowledge base supports the updated design, the replication ends. Otherwise, transitioning to a *recreation* replication type is necessary to field test the updated design. As a combination of *redesign* and *justification* (i.e. including the design and rigor cycles), this type of

replication study addresses all design theory components except for the purpose and scope and testable proposition.

Considering the example study, the demonstration and evaluation of the artifact took place in late 2013 and 2014, respectively. The study's screenshots indicate that the IS artifact had not been developed with smartphone users in mind. The mobile usage of the prototype may be quite cumbersome if not impossible. Recent advances in smartphone technologies and mobile app ecosystems would likely influence and change the previous results. For example, students may interact with smartphones differently than they do with desktop computers (e.g. due to mobile push notifications). Moreover, DP 1 (providing novel information to afford users disruptive ambiguity) would likely appear quite different if implemented on a mobile platform. Also, as users interact more frequently with their smartphones than with their desktop computers, the stimulus mechanisms used to pique interest could be quite different. For example, emotional social media campaigns may stimulate disruptive ambiguity better than information on the artifact platform. Accordingly, the existing artifact and design theory could be updated to the current state of web and mobile technologies, thus confirming, bounding, or contrasting previous findings. If an evaluation of the application environment is necessary, the study evolves into a *recreation* replication study.

Type 7: recreation

Is the design theory well justified and complete, and does it provide utility? Based on the concept of the progression framework (i.e. transitioning from one study type to another), the *recreation* type constitutes the most comprehensive form of replication. In a *recreation* study, researchers replicate all research activities related to relevance, rigor, and design. This study type validates the original design and re-examines it from different angles. If the original design withstands scrutiny, the overall replication is ultimately successful. However, a *recreation* replication can lead to a substantially different and overhauled design theory. This type of replication study can address any component of a design theory, as it attempts to replicate the entire artifact and all components of the design theory.

In addition to focus on organizational sensemaking theories, our example study draws heavily on focus group interviews to gain feedback to improve the artifact (the collaboration platform for organizational sustainability measures). The study's first theory-driven prototype may be necessary for academic publication; however, the current state of co-creative innovation management advises scholars to start by considering target group users (i.e. user-first approaches rather than theory-first endeavors). As the study begins theoretically, the users are presented with an initial prototype. The target groups' underlying needs and related desires are not analyzed until the first iteration is completed.

As design thinking suggests (Brown, 2008), considering the users may be a better approach, deeming the initial DP 1 not the most promising approach. The primary interest or need of a target group is not necessarily the provision of facts. Moreover, a completely different and likely latent need could be present, independent of the factual presentation. For instance, users may be more engaged in using web platforms for social interaction, such as joint team discussions and activities, than for discovering pure facts. Thus, a *recreation* replication study reiterates the entire DSR cycle. Researchers should address the design evaluation and ground their research process in the knowledge base to find an artifact that constitutes an improved design compared with the original study.

Type 8: meta-replication. The previously described replication types prescribe an active role for researchers in reaching a state where practitioners can adopt design theories and artifacts more easily. However, some artifacts and design theories have already reached a state where practitioners are well aware of these design theories and can readily implement them. The practical application and adaptation of artifacts can reveal design theory elements that could be relevant for developing solutions to other problems. Although these problems may be similar in some ways, they can still be very difficult. By observing how an artifact or design theory has been adopted in practice, researchers can learn about the problem class and related solution design. Similar to meta-studies or literature reviews, researchers can assume the seemingly passive role of analyzing the replication of artifacts and theories in their environment (e.g. how different users and organizations interact with a particular artifact). Thus, the distributed and implicit replication knowledge available in the artifact's environment can be captured and encoded. This replication type provides the opportunity to reflect on the existing design to elicit implications for other problem classes.

Similar to the example study and example described in the *update* study type section, the mobile context may also be assumed for *meta-replication*. Rather than actively updating the existing artifact with new technological means, researchers may simply observe the artifact's dissemination in new contexts. This observation would require that the artifact is readily available.

One may also consider IS that have adopted certain DPs prescribed by the case study. Similarities and differences may be identified through pure observation or through accompanying literature reviews (e.g. by conducting a forward citation search of the study and analyzing its academic dissemination). For example, by providing interactive communication features so that the system allows users to engage in an open and inclusive discussion in environmental sustainability transformation, the DP 3a may still be valid, yet more difficult to archive. Increasingly researched influences, such as bots, fake news, or a more politicized society, may illustrate that this DP remains relevant, but its current description could likewise become too vague to be implementable as is. Practitioners may have solved this problem by innovative filtering mechanisms that are relevant to all similar instances of this problem. Hence, certain adaptations to technological changes may become apparent and could be gathered, summarized, and synthesized to improve the original design theory.

Discussion

This study proposes replication as a valuable means for DSR to increase confidence in design theories. Confidence is an important prerequisite for reuse and knowledge accumulation, which are practices that are severely underrepresented in DSR (Kruse and Seidel, 2017; Vom Brocke et al., 2020). A substantial knowledge gap exists regarding how replication fits within the research paradigm of DSR and how it should be conducted. To contribute to closing this gap, we developed eight DSR replication study types, building upon a transfer of replication logic from descriptive inquiry to the context of prescriptive inquiry (March and Smith, 1995).

Regarding the fit of replication within the paradigm of DSR, we contributed to IS research by conceptualizing the differences between descriptive and prescriptive research regarding the application of replication research. Prescriptive knowledge cannot be refuted because utility is a time- and context-dependent value. A new design theory may be more useful than an existing one but does not thereby refute the existing one. Hence, the replication of design theories is fundamentally different from the replication of other types of theories (Dennis and Valacich, 2014; Gregor, 2006), rendering two disjointed subcategories of replication research: replication of prescriptive knowledge (utility-oriented) and replication of descriptive knowledge (truth-oriented). This proposition is supported by the structural differences between the respective replication study types. Based on Brendel et al. (2020), replication study types can be distinguished along the dimensions of theory, method, and context and how similar or dissimilar the replication study is to the original study. Because DSR does not employ an exact method, but rather conducts an iterative research process, consisting of many different methods, it is nearly impossible to exactly replicate a DSR study's research appraoch. Similarly, the artifact's application environment is constantly changing and evolving (Olbrich et al., 2017), making it impractical to replicate a previous environment entirely.

This study proposes replicating design theories in a similar environment via a progressive research approach. Researchers can adapt their research processes according to gathered insights, extending the research process and the implications for the design theory. Hence, the subcategories of replication study types are different in terms of their replication activities and their expected replication outcomes (i.e. confirming or changing the components of the original design theory).

Critics have complained that the DSR community has failed to reuse design theories and accumulate knowledge (Vom Brocke et al., 2017, 2020). This study's conceptualization of multiple replication research types in DSR is designed to increase confidence in design theories. The eight DSR replication study types and their progressive connecting structure contribute to this process in two ways. First, replication is a form of reuse because a replication study examines, applies, and extends existing design theories. Second, replicating design theories tests their validity, which results in an increase or decrease in the assumed confidence level of a design theory. Furthermore, if a replication attests that a design theory prescribes an artifact's design such that it sufficiently solves a problem, then the confidence of scholars in this design theory increases. A high level of confidence indicates to researchers that a design theory can be reused to derive new design theories. Hence, this study contributes to DSR by providing a means to increase confidence in design theories, which is a vital prerequisite for reusing design theories and accumulating knowledge in the DSR community.

The introduction of the replication concept to DSR also contributes to practice. Increasing confidence in design theories enables researchers to formulate precisely when and how a design theory is applicable and what can be expected from an artifact. After a design theory reaches a sufficient level of confidence, practitioners can use it as a blueprint for real-world artifact development more safely.

This study has some limitations, which suggest potential avenues for future research. First, the review of the literature regarding knowledge on replication in DSR (see Appendix) identified research on some topics related to replication research (e.g. reuse, projectability). However, these studies lacked an overarching integration into replication research. Building on the progression framework presented in this study, future research may draw a more complete picture of what has been directly and indirectly conceptualized regarding replication in DSR and may reduce gaps or blind spots that may still exist in the field.

Second, replication is based on the concept of falsification (i.e. a theory can either be true or useful in the case of DSR or false or useless in the case of DSR). A more nuanced view suggests "confidence" as a term to describe an attribute of theoretical statements or propositions. From this viewpoint, a theory can evolve from (1) being proposed with little or no supporting evidence via (2) many tests and re-examinations to (3) a greater maturity, perhaps with more precise wording. The theory may or may not change over time, but the level of confidence can increase with each supporting test and decrease with contradictory results. Future research should conceptualize how confidence fits within other DSR concepts, such as maturity and abstraction (Gregor and Hevner, 2013).

Third, this study synthesized the proposed replication study types as a progression framework. This framework illustrates the interrelations and potential transitions between different replication study types. This approach has great potential, within a few years, for gathering knowledge about how and when to transit between these phases (i.e. how to interpret gathered evidence to determine an appropriate transition adequately).

Fourth, this study proposes DSR replication study types based on literature, logical reasoning, and thought experiments. It introduces a comprehensive set of study types that can support researchers in their replication endeavors. To strengthen the foundation of these study types further—and also to foster their implementation—we would like to call future research to investigate the literature systematically. Such investigation could identify studies that can (at least partly) be related to these replication study types. For example, some studies might have discussed a design theory by comparing it to existing explanatory theories. Such studies might raise contradictions, which would match with the *justification* replication study type. Learning from this study could provide first insights to formulate guidelines for conducting *justification* replications.

Conclusion

This article addresses the issue of missing artifact and design theory reuse and the resulting lack of knowledge accumulation in DSR. It provides researchers with an initial understanding of the role of replication research in DSR to increase confidence, which is a key prerequisite for reuse and knowledge accumulation. It proposes eight DSR replication study types and connects these study types in a progressive structural framework, enabling researchers either to develop an improved version of the replicated design theory or to increase confidence in the original version. The set of replication study types and the progression framework can guide editors and reviewers in identifying and evaluating replication studies for DSR projects.

As with similar conceptual articles, the replication research questions and subsequent replication study types serve as a starting point and are not intended to be definitive. The purpose is rather to illustrate how research might engage replication in DSR, providing an initial framework. Hence, researchers should take these ideas about issues and potential solutions regarding replication in DSR as motivation to begin implementing replication research, either by following this framework or deliberately deviating from it.

The time for replication research in DSR has come, and we anticipate greater interest in DSR replication studies. This article aims to provide interested scholars with an initial reference and "food for thought" for their replicative endeavors.

Declaration of conflicting interests

Funding

ORCID iDs

Alfred Benedikt Brendel (iD) https://orcid.org/0000-0002-6511 -221X

Tim-Benjamin Lembcke (iD) https://orcid.org/0000-0003-3092-5277

References

Agerfalk PJ (2010) Getting pragmatic. *European Journal of Information Systems* 15(1): 251–256.

Anderson SF and Maxwell SE (2016) There's more than one way to conduct a replication study: Beyond statistical significance. *Psychological Methods* 21(1): 1–12.

Arnott D (2006) Cognitive biases and decision support systems development: A design science approach. *Information Systems Journal* 16: 55–78.

Arnott D and Pervan G (2012) Design science in decision support systems research: An assessment using the Hevner, March, Park, and Ram guidelines. *Journal of the Association for Information Systems* 13(11): 923–949.

Baker M (2015) Over half of psychology studies fail reproducibility test. *Nature—News*. Available at: https://www.nature.com/news/over-half-of-psychology-studies-fail-reproducibility-test-1.18248 (accessed 2 April 2021).

Baker M (2016) 1,500 scientists lift the lid on reproducibility. *Nature* 533(7604): 452–454.

Banker RD and Kauffman RJ (2004) 50th anniversary article: The evolution of research on information systems: A fiftieth-year survey of the literature in management science. *Management Science* 50(3): 281–298.

Baskerville R, Baiyere A, Gregor S, et al. (2018) Design science research contributions: Finding a balance between artifact and theory. *Journal of the Association for Information Systems* 19(5): 358–376.

Baskerville R, Kaul M and Storey VC (2017) Establishing reliability in design science research. In: *Proceedings of international conference on information systems*, Seoul, South Korea, December 2017, p. 18. Association for Information Systems.

Beck R, Weber S and Gregory RW (2013) Theory-generating design science research. *Information Systems Frontiers* 15(4): 637–651.

Bonett DG (2012) Replication-extension studies. *Current Directions in Psychological Science* 21(6): 409–412.

Brendel AB, Greulich RS, Niederman F, et al. (2020) Towards a greater diversity of replication studies. *AIS Transactions on Replication Research* 6(1): 20.

Brendel AB, Zapadka P and Kolbe LM (2018) Design science research in green IS: Analyzing the past to guide future research. In: *Proceedings of European conference on information systems*, Portsmouth, United Kingdom, June 2018, pp. 1–18 Association for Information Systems.

Brown T (2008) Design thinking. *Harvard Business Review* 86(6): 84–92.

Compeau D, Marcolin B, Kelley H, et al. (2012) Research commentary-generalizability of information systems research using student subjects: A reflection on our practices and recommendations for future research. *Information Systems Research* 23(4): 1093–1109.

Cooper HM (1988) Organizing knowledge synthesis: A taxonomy of literature reviews. *Knowledge in Society* 1(1): 104–126.

Deci EL, Ryan RM, Vallerand RJ, et al. (1991) Motivation and education: The self-determination perspective. *Educational Psychologist* 26(3–4): 325–346.

Dennis AR and Valacich JS (2014) A replication manifesto. *AIS Transactions on Replication Research* 1(1): 1–5.

Goldkuhl G (2012) Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems* 21: 135–146.

Gómez OS, Juristo N and Vegas S (2010) Replications types in experimental disciplines. In: *Proceedings of the 20 ACM-IEEE international symposium on empirical software engineering and measurement—ESEM '10*, Bolzano-Bozen, Italy, September.

Gómez OS, Juristo N and Vegas S (2014) Understanding replication of experiments in software engineering: A classification. *Information and Software Technology* 56(8): 1033–1048.

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

Gregor S and Hevner AR (2013) Positioning and presenting design science research for maximum impact. *MIS Quarterly* 37(2): 337–355.

Gregor S and Jones D (2007) The anatomy of a design theory. *Journal of the Association for Information Systems* 8(5): 312–334.

Gregory RW and Muntermann J (2014) Heuristic theorizing: Proactively generating design theories. *Information Systems Research* 25(3): 639–653.

Greulich RS and Brendel AB (2019) Towards a greater appeal of replication research by learning from the field of neuroscience. In: *Proceedings of Americas conference on information systems*, Cancun, Mexico, August 2019, p. 10. Association for Information Systems.

Herwix A and Rosenkranz C (2018) Making sense of design science in information systems research: Insights from a systematic literature review. In: *International conference on design science research in information systems and technology*, Chennai, India, 3 June.

Hevner AR (2007) A three cycle view of design science research. *Scandinavian Journal of Information Systems* 19(2): 87–92.

Hevner AR, March ST, Park J, et al. (2004) Design science in information systems research. *MIS Quarterly* 28(1): 75–105.

Iivari J (2015) Distinguishing and contrasting two strategies for design science research. *European Journal of Information Systems* 24(1): 107–115.

Kruse LC and Seidel S (2017) Tensions in design principle formulation and reuse. In: *Proceedings of international conference on design science research in information systems and*

technology, Karlsruhe, Germany, 30 May–1 June, pp. 1–10. Karslruhe: Karlsruher Institut für Technologie.

Kuechler W and Vaishnavi V (2012) A framework for theory development in design science research: Multiple perspectives. *Journal of the Association for Information Systems* 13(6): 395–423.

Lee A and Nickerson JV (2010) Theory as a case of design: Lessons for design from the philosophy of science. In: *Proceedings of Hawaii international conference on system sciences*, Honolulu, HI, 5–8 January 2010, pp. 1–10. New York: IEEE.

Lee AS and Hubona GS (2009) A scientific basis for rigor in information systems research. *MIS Quarterly* 33(2): 237–262.

Lee JS, Pries-Heje J and Baskerville R (2011)Theorizing in Design Science Research. In: Jain H, Sinha AP and Vitharana P (eds) *Service-Oriented Perspectives in Design Science Research. DESRIST 2011*. Lecture Notes in Computer Science, vol 6629, Berlin, Heidelberg: Springer.

Leukel J, Mueller M and Sugumaran V (2014) The state of design science research within the BISE community: An empirical investigation. In: *Proceedings of international conference on information systems*, Auckland, New Zealand, December 2014, pp. 1–15. Association for Information Systems.

Levy Y and Ellis TTJ (2006) A systems approach to conduct an effective literature review in support of information systems research. *Informing Science Journal* 9: 181–211.

March ST and Smith GF (1995) Design and natural science research on information technology. *Decision Support Systems* 15(4): 251–266.

Melville NP (2010) Information systems innovation for environmental sustainability. *MIS Quarterly* 34(1): 1–21.

National Academies of Sciences, Engineering, and Medicine (2019) *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press.

Niederman F and March S (2014) Reflections on replications. *Transactions on Replication Research* 1: 1–16.

Olbrich S, Frank U, Gregor S, et al. (2017) On the merits and limits of replication and negation for IS research. *Transactions on Replication Research* 3: 1–19.

Österle H and Otto B (2010) Consortium research. *Business & Information Systems Engineering* 2(5): 283–293.

Peffers K, Tuunanen T, Rothenberger MA, et al. (2007) A design science research methodology for information systems research. *Journal of Management Information Systems* 24(3): 45–77.

Peng RD (2011) Reproducible research in computational science. *Science* 334(6060): 1226–1227.

Popper KR (1959) *The Logic of Scientific Discovery*. London: Hutchinson & Co.

Popper KR (1963) *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge and Kegan Paul.

Ryan RM and Deci EL (2004) Autonomy is no illusion: Self-determination theory and the empirical study of authenticity, awareness, and will. In: Greenberg J, Koole SL and Pyszczynski T (eds) *Handbook of Experimental Existential Psychology*. New York: Guilford Press, pp. 455–485.

Salovaara A and Merikivi J (2015) IS research progress would benefit from increased falsification of existing theories. In: *Proceedings of European conference on information systems*, Muenster, Germany, May 2015, p. 16. Association for Information Systems.

Schmidt S (2009) Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology* 13(2): 90–100.

Seidel S, Chandra Kruse L, Székely N, et al. (2018) Design principles for sensemaking support systems in environmental sustainability transformations. *European Journal of Information Systems* 27(2): 221–247.

Sein MK, Henfridsson O, Purao S, et al. (2011) Action design research. *MIS Quarterly* 35(1): 37–56.

Venable J (2006) The role of theory and theorising in design science research. In: *Proceedings of international conference on design science research in information systems and technology*, Claremont, California, USA, June 2006, pp. 24–35. Association for Information Systems.

Vom Brocke J, Hevner A, Maedche A, et al. (2017) Call for papers special issue JAIS: Accumulation and evolution of knowledge in design science research. *Journal of the Association for Information Systems* 21: Article 9.

Vom Brocke J, Simons A, Niehaves B, et al. (2009) Reconstructing the giant: On the importance of rigour in documenting the literature search process. In: *Proceedings of European conference on information systems*, Verona, Italy, June 2009, pp. 2206–2217. Association for Information Systems.

Vom Brocke J, Winter R, Hevner A, et al. (2020) Accumulation and evolution of design knowledge in design science research—A journey through time and space. *Journal of the Association for Information Systems* 21: 520–544.

William J and Choong C (1999) Replication of measures of information systems research: The case of IS SERVQUAL. *Decision Sciences* 30(3): 893–899.

Winter R (2008) Design science research in Europe. *European Journal of Information Systems* 17: 470–475.

Wynn D and Williams CK (2012) Principles for conducting critical realist case study research in information systems. *MIS Quarterly* 36(3): 787–810.

Author biographies

Alfred Benedikt Brendel is an associate professor of Business Informatics, esp. Intelligent Systems and Services, at the Technische Universität Dresden, Germany. His research is concerned with the design of information systems in the domains of healthcare, transportation, and digital workplace. Alfred's research has been published in leading IS journals, such as *Technological Forecasting and Social Change, Business & Information System Engineering*, and *Information Systems Frontiers*.

Tim-Benjamin Lembcke is a research associate at the University of Göttingen and the current head of the Smart Mobility Research Group. He conducts research on digital nudging and the related ethical concerns. His research has been published in various outlets, including *Transportation Research Part D: Transport and Environment and the proceedings of the International Conference on Information Systems*.

Jan Muntermann is a full professor of electronic finance and digital markets at the University of Göttingen, Germany. His research interests include (big) data analytics and managerial decision support, digital business strategy development and execution, and the conceptual and methodological foundations of theory development in information systems research. He has published in journals, such as *Information Systems Research, Decision Support Systems, and the European Journal of Information Systems*.

Lutz M Kolbe is a full professor and Chair of Information Management at the University of Göttingen, Germany, since 2007. At the Technical University Freiberg, Germany, he earned his doctorate in Business Administration. Subsequently, he led Deutsche Bank's IT think tanks in Frankfurt, New York, and San Francisco. In 2002, he moved on to the Institute of Information Systems at the University of St. Gallen, Switzerland, as head of several competence centers. His research has been published in leading international journals, including, *Information Systems Journal, European Journal of Information Systems, and MIS Quarterly Executive*.

Appendix

The search for literature on replication research in DSR followed the process outlined by Vom Brocke et al. (2009). As a first step, we selected relevant publication outlets for our literature search, restricting our list to top-tier journals to ensure relevance, rigor, and impact (Levy and Ellis, 2006). Hence, we focused on articles published in the eight top-ranked IS journals (i.e. the "basket of eight"):

1. *Information Systems Research* (ISR);
2. *Management Information Systems Quarterly* (MISQ);
3. *Journal of Management Information Systems* (JMIS);
4. *Journal of the Association for Information Systems* (JAIS);
5. *Journal of Information Technology* (JIT);
6. *Information Systems Journal* (ISJ);
7. *The Journal of Strategic Information Systems* (JSIS);
8. *European Journal on Information Systems* (EJIS).

Our search also included *AIS Transactions on Replication Research*, a journal focused on replication research that was first published in 2015.

To complement articles sourced from journals, we decided to include articles from conference proceedings and added publications from the following conferences to our search:

- International Conference on Information Systems (ICIS);
- European Conference on Information Systems (ECIS);
- Pacific Asia Conference on Information Systems (PACIS);
- Americas Conference on Information Systems (AMCIS).

To locate studies, we used the Web of Science, AIS Electronic Library databases, and the search comprised the full text. The search also included the websites for these journals and conferences if the outlet was not already included in the databases. We applied the following query in our literature search:

"Replication" AND ("Design Science" OR "Design Research")

Two scholars conducted the literature search in September and October 2019. To determine the relevance of the identified articles, we analyzed the articles in several steps. First, we briefly scanned the titles and abstracts and removed irrelevant articles. For example, the term "replication" also describes the simulation of behavior ("replicating human behavior") in articles not relevant to this study. Second, we removed any articles that did not conduct replication research in DSR or did not address the topic of replication research in DSR.

We did include studies that had brief sections on this topic (as part of their discussion or conclusion). For example, Arnott (2006) concluded in his study that the presented results may not be generalizable to other projects and needed further research, possibly following a "replication logic" (p. 73).

Overall, we identified six papers that at least partly address the topic of replication in DSR (Table 4). The manner in which the topic was approached varied among the papers, with specific subtopics and formats as follows: challenges in replicating DSR study results (Arnott, 2006); general calls for replication in DSR (Dennis and Valacich, 2014; Niederman and March, 2014); challenges of uncertainty regarding results of critical realist case studies followed by mention of DSR (Wynn and Williams, 2012); replication in DSR as part of overall rigor in IS research (Lee and Hubona, 2009); a side-topic within the context of reliability in DSR (Baskerville et al., 2017).

Table 4. Summary of literature search.

Source type	Number of articles found	
Journals	Total	Filtered
Information Systems Research (ISR)	5	0
Management Information Systems Quarterly (MISQ)	19	2
Journal of Management Information Systems (JMIS)	8	0
Journal of the Association for Information Systems (JAIS)	27	0
Journal of Information Technology (JIT)	2	0
Information Systems Journal (ISJ)	4	1
The Journal of Strategic Information Systems (JSIS)	4	0
European Journal on Information Systems (EJIS)	6	0
AIS Transactions on Replication Research (TRRJ)	2	2
Subtotal	77	5
Conferences		
International Conference on Information Systems (ICIS)	87	1
European Conference on Information Systems (ECIS)	82	0
Pacific Asia Conference on Information Systems (PACIS)	33	0
Americas Conference on Information Systems (AMCIS)	78	0
Subtotal	280	1
Total	357	6