Information Systems Design Knowledge for Sustainable Development Along a Social-Technical Continuum

Dissertation

zur Erlangung des akademischen Grades
Dr.-Ing.

eingereicht an der
Mathematisch-Naturwissenschaftlich-Technischen Fakultät
der Universität Augsburg

von

Sebastian Heger, M.Sc.

Augsburg, Juni 2020

Universität
Augsburg
University

Erstgutachter:	Prof. Dr. Henner Gimpel
Zweitgutachter:	Prof. Dr. Axel Tuma
Datum der Disputation:	28. Juli 2020

"The greatest threat to our planet is the belief that someone else will save it."

Robert Swan

Abstract

Achieving a more sustainable lifestyle is one of the most important challenges of the 21st century. Sustainable development, therefore, is one of the key objectives also for industrial nations such as Germany. Sustainable development is defined as the "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (WCED 1987, Chapter 2). This development, however, is not only based on environmental dimensions but also on economic and social dimensions. These are known as the three pillars of sustainability.

According to the Global Sustainable Development Report (Messerli et al. 2019), there are four *levers of transformation* being most important to achieve sustainable development, one of which is science and technology. Information systems (IS) research investigates the application of information technology in organizational settings (Hevner et al. 2004). IS research considers itself as a socio-technical discipline (Briggs et al. 2010), which should investigate IS-related issues along a social-technical continuum (Sarker et al. 2019). In recent times, the field is discussing more intensively the application of IS for sustainability purposes (e.g., Melville 2010; Watson et al. 2010; Henkel and Kranz 2018). However, IS scholars claim for more work to be done regarding sustainable development (Seidel et al. 2017; Gholami et al. 2016; Parmiggiani and Monteiro 2018) and to develop knowledge along the social-technical continuum (Sarker et al. 2019).

For this reason, the aim of this dissertation is to develop knowledge for IS to achieve sustainable development from different socio-technical perspectives. Therefore, this work applies pluralistic methodological approaches (qualitative methods, e.g., semi-structured interviews, and quantitative methods, e.g., data collection from a field study or online experiments). The theoretical contribution of this dissertation expands existing design knowledge in the field of IS for sustainability. Following Gregor and Hevner (2013), design knowledge can be grouped into two types: descriptive and prescriptive knowledge. On the one hand, descriptive knowledge contributes to the knowledge about natural phenomena and the sense-making relationships between phenomena (what). Descriptive knowledge provides the scientific base for the world we live in. On the other hand, prescriptive knowledge contributes to the knowledge about man-made artifacts to improve our world (how). Design theories are an abstract and coherent body of knowledge, which include both types of knowledge (Gregor and Hevner 2013). Overall, this work takes three different perspectives to develop both types of knowledge: the socio-technical perspective, the sociocentric perspective, and the technocentric perspective.

First, to contribute to the socio-technical perspective, this thesis presents new approaches to analyze and improve individual driving behavior in individual road traffic. Personal traffic accounts for approximately 11% of carbon dioxide (CO_2) emissions globally (Andor et al. 2020). Environmental driving behavior, therein, has a significant impact on the fuel consumption of vehicles (Lárusdóttir and Ulfarsson 2014), and reductions of fuel consumption up to 30% are possible. One promising approach to improve individual driving behavior is the application of eco-feedback to the driver. Therefore, this thesis presents prescriptive knowledge in the form of design artifacts for mobile eco-driving feedback information systems (EDFIS), which results from justificatory knowledge, a prototypical instantiation of a mobile EDFIS, and its application within a field study. The results indicate that eco-feedback affects environmentally friendly driving behavior. However, eco-feedback seems not to address all aspects of eco-driving behavior. Nowadays, the Internet-of-Things (IoT) has reached cars, and ever more embedded sensors allow for rich data analysis of individual driving behavior. Therefore, this thesis presents a factor model describing IoT-measured individual driving behavior to handle the resulting amount of IoT-data and analyzes the effect of eco-feedback on individual driving behavior.

Second, to contribute to the sociocentric perspective, this thesis presents design knowledge on how to design social media platforms in order to counteract so-called fake news. Such misleading information – which has always been a problem but has recently found nutritious ground in social media – poses a major threat. In recent times, fake news is spread to impact political decision-making and elections. However, fake news is not only a challenge for politics but also for ecologically and economically sustainable development. For instance, fake news is spread to influence financial markets in order to affect the financial value of stocks and options (Maasberg et al. 2018). Furthermore, fake news is a popular means to undermine climate change, which in reality requires significant changes in the individual and collective behavior of people (van der Linden et al. 2017). This thesis presents two new approaches to empower people individually to recognize and report misleading information to counteract the spread of fake news. Firstly, related articles are an appropriate tool to improve social media users' ability to recognize fake news as such. Especially, the application of controversial related articles shows the best results. Secondly, social norm messages are an appropriate tool to encourage social media users to report fake news. This allows social media providers to effectively identify and remove misleading and deceptive information from their platforms.

To contribute to the third and last, technocentric perspective, this thesis presents a new approach to analyze and design lean, digitally supported value creation processes in companies and organizations. In the light of ever more connected entities and the diffusion of digital technologies, innovative information flows enable various potentials in value creation processes and allow minimizing waste. For instance, rich information availability allows production with fewer resources or the more accurate usage of perishable materials. The presented Value Stream Modeling and Notation (VSMN) constitutes a domain-specific modeling language, which supports designers of economically sustainable value creation processes in times of digitalization. For evaluation purposes, we present, among others, the results of the application of VSMN while designing lean material logistics processes in hospitals. The results also constitute prescriptive knowledge for material logistic processes in hospitals, which are more efficient due to the targeted use of digital technologies.

To sum up, this dissertation presents design knowledge – including both types, descriptive and prescriptive knowledge – and contributes to the knowledge base about IS for sustainability (Gholami et al. 2016; Seidel et al. 2017; Parmiggiani and Monteiro 2018) and sustainable development (Messerli et al. 2019). In addition, the thesis develops design knowledge along the social-technical continuum (Sarker et al. 2019). Using methodological pluralistic (qualitative and quantitative) approaches allows an extension of the existing knowledge from different perspectives (Venkatesh et al. 2013). Overall, the results of this thesis contribute to overcoming one of the greatest challenges of mankind and the greatest challenge of the 21st century.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA .... *Analysis of variance*
AVE .... *Average variance extracted*
BPMN .... *Business Process Model and Notation 2.0*
CFA .... *Confirmatory factor analysis*
CO_2 .... *Carbon dioxide*
CR .... *Composite reliability*
DO .... *Design objective*
DP .... *Design principle*
DSML .... *Domain-specific modeling language*
EDFIS .... *Eco-driving feedback information system*
EFA .... *Exploratory factor analysis*
GHG .... *Greenhouse gas*
IoT .... *Internet-of-Things*
IS .... *Information system*
IT .... *Information technology*
MaMa .... *Materials management*
n.d. .... *No date*
OBD-II .... *On-board diagnostic II*
RPM .... *Revolutions per minute*
SDG .... *Sustainability Development Goals*
SN .... *Social norm*
UDI .... *Unique device identifier*
UML .... *Unified Modeling Language*
VDI .... *Association of German Engineers*
VSM .... *Value stream method*
VSMN .... *Value Stream Modeling and Notation*

1. Introduction

1.1. Motivation

Achieving a more sustainable lifestyle is one of the most important challenges of the 21st century. Today, however, we are failing to achieve this goal. Most critically, the consumption of natural resources exceeds by far the regenerative capacity of our nature. According to the FootPrint Network (2020), earth overshoot day marks the date when humanity consumed more resources than naturally available for one year. In 2019, the date was July 29th. However, when all countries may consume as many resources as Germany, this day would shift to May 3rd.

As a result, harmful effects are already visible today. Global warming due to greenhouse gas (GHG) emissions already increased global average temperature about one degree since the beginning of industrialization. And scientists' forecasts promise drastic consequences such as rising sea levels, prolonged droughts, ocean acidification, and loss of biodiversity when a further rise of three to four degrees by 2100 takes place (IPCC 2014). Cleary, these environmental damages are affecting humans, too, and will even become more dramatic. For instance, poor access to drinking water, food, and increasing diseases (e.g., dengue, malaria) (IPCC 2014). But also collapsing economic sectors are a severe effect, resulting in poverty, unemployment, and so-called environmental refugees (Climate Justice Network n.d.).

To counteract these hazards, also industrial nations such as Germany must contribute to sustainable development. Sustainable development is defined as the “development that meets the needs of the present without compromising the ability of future generations to meet their own needs” (WCED 1987, Chapter 2). What is more, sustainable development is not only based on environmental dimensions but also on an economic and social aspect (Elkington 1994; Sarkis et al. 2013). Thus, sustainable development is a complex task to achieve.

Since 2015 the United Nations provide the Sustainability Development Goals (SDG) to address sustainable development globally (United Nations 2015). The SDG include seventeen different goals regarding economic, environmental, and societal dimensions, which should be achieved by 2030. The SDG not only suggest that sustainable development must address all three pillars but also that the dimensions are, in part, closely related to each other.

Based on the SDG, in 2019, the United Nations introduced different entry points, which are deemed most important to enable sustainable development (Messerli et al. 2019). For instance, the entry point *energy decarbonization and universal access* addresses sustainable development by calling for full access to clean and green energy but also improving energy efficiency. By the year 2040, energy demand may increase by another 25% due to rising incomes and a larger population (Messerli et al. 2019). However, the additional demand will be much higher if no further work is done on energy efficiency measures. Energy efficiency is an effective way to reduce the consumption of fossil energy sources and, as a consequence, to decrease GHG emissions. Besides, energy efficiency also results in an economic advantage in the long term (Messerli et al. 2019). Comparable, *sustainable and just economies*, as another entry point, address the fact that large parts of economies are not sustainable today in a way that has negative impacts on the environment and society. Therefore, it is mandatory to achieve long-term sustainability pathways and to decouple economic growth from resource overuse (Messerli et al. 2019). While some argue that decoupling economic growth from resources is already happening to some extend (e.g., McAfee 2019), others show that there is no evidence for a decoupling today (e.g., Parrique et al. 2019). As a last example, the entry point *human well-being and capabilities* addresses improvements in quality of life, such as affordable healthcare and education. These are vital aspects of sustainable development, which require to improve human capabilities – such as knowledge and (digital) competencies – in a way that people can foster change on an environmental, economic, and social level (Messerli et al. 2019).

Messerli et al. (2019) also identify four levers of transformation to address these various entry points for sustainable development, one of which is *science and technology*. Science has to find ways and means to achieve sustainable development, which is a challenge that involves many academic disciplines (e.g., Watson et al. 2010; Gholami et al. 2016). Technology, on the other hand, is a fundamental aspect of achieving sustainable objectives. That is, science has to develop and improve technologies but also find ways to apply technologies in an appropriate way (Messerli et al. 2019). One field of research, dealing with the target-oriented application of technology and, especially, digital technologies for sustainability is information systems (IS) research (Melville 2010; Gholami et al. 2016; Watson et al. 2010; Walsham 2017).

1.2. The Role of Information Systems in Sustainable Development

IS research investigates the interaction between people, organizations, and technology (Hevner et al. 2004). From a fundamental viewpoint, IS has long been a discipline “at the crossroads of many social and technical disciplines” (Briggs et al. 2010, p. 14). IS contributions add to the understanding of both kinds of disciplines concerning information technology (IT) playing a central role. Hence, IS is a socio-technical discipline (Briggs et al. 2010; Sarker et al. 2019). A socio-technical system itself consists of two independent but interacting systems. The technical system includes "processes, tasks, and technology needed to transform inputs to outputs" (Bostrom and Heinen 1977, p. 17). The social system is more about "the attributes of people (e.g., attitudes, skills, values), the relationships among people, reward systems, and authority structures" (Bostrom and Heinen 1977, p. 17). IS research must not only take into account both perspectives separately but rather has to develop knowledge about the interaction of both perspectives (Sarker et al. 2019). Yet, Sarker et al. (2019) argue that over the last two decades, IS research has apparently lost its focus on a socio-technical perspective. There is a dominant shift towards a social perspective (where technology is only a context), and, into the bargain, a trend is discernible towards a strengthening of the technical perspective (leaving social considerations behind). For this reason, IS researchers call for action to develop IS knowledge from multiple theoretical and methodological approaches along a social-technical continuum, as described in Figure 1 (Sarker et al. 2019).

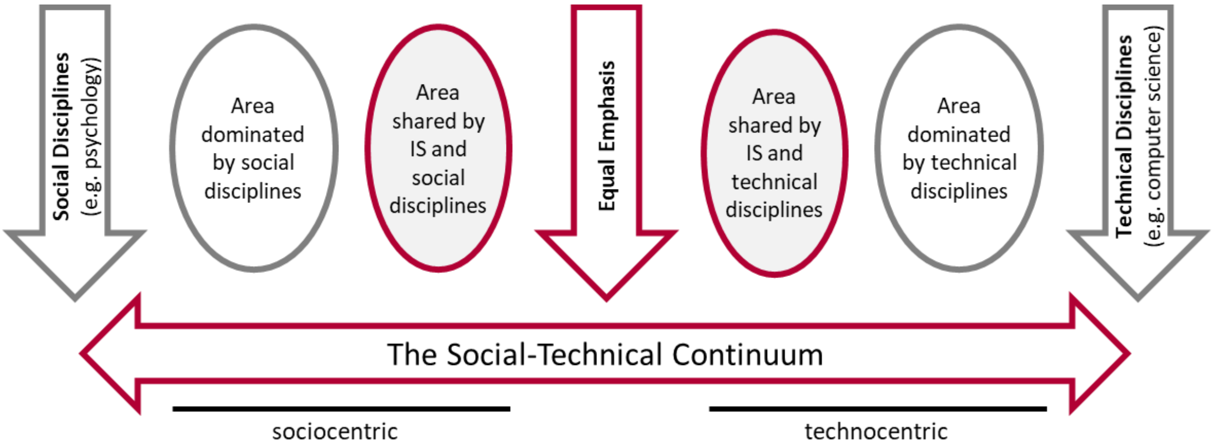


Figure 1: The social-technical continuum following Sarker et al. (2019)

The structural nature of IS knowledge consists of five interrelated types of theories, one of which is design theory (Gregor 2006). A design theory "is about the principles of form and function, methods, and justificatory theoretical knowledge that are used in the development of IS" (Gregor 2006, p. 628). Following Gregor and Hevner (2013), design knowledge can be grouped into two types: descriptive (about what) and prescriptive knowledge (about how), which is together encapsulated in design theories.

Over the last decades, IS researchers published both types of knowledge to design IS to counteract the adverse effects of human behavior to achieve environmental sustainability. Also, IS arguably has substantial impacts on economic development and, thus, researchers also investigated how to enable sustainable economic development (Watson et al. 2010). A major research stream in IS dealing with both environmental and economic sustainability is called *IS for Environmental Sustainability*, which is defined as "IS-enabled practices and processes improving environmental and economic performance" (Melville 2010, p. 8). The research stream focuses on emerging and diffusing environmentally sustainable practices through the application of so-called green IS by individuals, groups, organizations, or societies (Henkel et al. 2017; Henkel and Kranz 2018; Watson et al. 2010). Green IS are defined as a "cooperating set of people, processes, software, and information technologies to support individual, organizational, or societal goals" (Kranz et al. 2015, p. 8). Green IS and IT have both been leveraged to support sustainable development (Kranz et al. 2015).

For this reason, the term green IS often is closely related to the concept of green IT (Parmiggiani and Monteiro 2018; Loeser 2013). However, there is a clear distinction (e.g., Sarkis et al. 2013; Kranz et al. 2015). Green IT refers to hardware and infrastructures that can be better managed and designed from an environmental perspective. It aims to reduce the negative environmental impacts of IT (Sarkis et al. 2013; Henkel and Kranz 2018). Research on green IT belongs to the technical end of the social-technical continuum. An objective of Green IS is its support of pro-environmental behaviors and actions (Henkel and Kranz 2018). Therefore, it can be argued that green IS research focuses on the equal emphasis and shared areas along the social-technical continuum.

Over the past decade, IS research has addressed the purposeful application of IS for sustainability. However, many scientists recently repeated that IS research still does not sufficiently investigate the issue of sustainable development (e.g., Gholami et al. 2016; Seidel et al. 2017; Parmiggiani and Monteiro 2018).

1.3. Aim and Outline of this Thesis

IS scholars emphasize that the IS discipline should redirect its focus on socio-technical systems (Sarker et al. 2019; Briggs et al. 2010). At the same time, IS scholars also call for increased work on IS for sustainability (Melville 2010; Watson et al. 2010; Gholami et al. 2016; Seidel et al. 2017). For these reasons, the doctoral thesis at hand aims at contributing to green IS design for sustainable development. The thesis also takes up the argumentation that IS knowledge is required across the social-technical continuum. Therefore, the thesis develops IS knowledge for sustainability from three perspectives following Sarker et al. (2019): a socio-technical perspective (equal emphasis of the social and technical system), a sociocentric perspective (social imperative, the social aligning the technical system), and a technocentric perspective (technical imperative, the technical system enabling the social system).

First, this thesis takes a socio-technical perspective. Design science, therefore, draws on theories from natural or social sciences to derive efficient designs. Hence, social considerations are included in technological artifacts (Sarker et al. 2019). One research discussion within green IS to foster pro-environmental behavior of individuals is the application of IS to increase energy efficiency in the sense of decreasing carbon dioxide (CO_2) emissions (Watson et al. 2010). As one of the main polluters, traffic accounts for approximately 20% of CO_2 emissions in Germany (UBA 2017). Environmental driving behavior has a significant impact on the fuel consumption of vehicles (Lárusdóttir and Ulfarsson 2014), and reductions of fuel consumption up to 30% are possible (Sivak and Schoettle 2012). Different studies investigated the effect of eco-driving feedback systems and found promising results of savings up to 7%. Although various studies have already investigated the effect of eco-feedback information systems (EDFIS), there is prescriptive knowledge missing on the design of EDFIS.

For this reason, this thesis presents design knowledge in the form of design artifacts for EDFIS, which results from the literature, a prototypical instantiation, and a field study. However, driving behavior is complex, and it is not yet clear which aspects are addressed by eco-driving feedback. Nowadays, the Internet-of-Things (IoT) has reached many cars, and ever more embedded sensors allow for rich data analysis of individual driving behavior. On these grounds, the doctoral thesis at hand presents a factor model describing IoT-measured individual driving behavior to handle the resulting amount of IoT-data. Further, it analyzes the effect of eco-feedback on individual driving behavior.

Second, this thesis takes a sociocentric perspective. The aim is to generate knowledge on how social structures and processes should influence the design of technologies. Concerning sustainable development, people have to grasp the challenges and take appropriate countermeasures. However, scientific findings and appeals are counteracted by non-environmental interest groups in that false information is deliberately disseminated (van der Linden et al. 2017). A threat, which gains increasing importance by the fact that social media becomes more and more popular as a platform for news consumption (Gottfried and Shearer 2016; Shu et al. 2017). Social media is a breeding ground for so-called fake news being news that "are intentionally and verifiably false" (Allcott and Gentzkow 2017). In recent years, scientists have gained a better understanding of why people respond to fake news (Moravec et al. 2019). Besides, they have also studied source ratings as an appropriate countermeasure to empower users to better recognize fake news as such (Kim et al. 2019; Kim and Dennis 2019). However, less has been done to support users when the source is somewhat unknown. Theory suggests that providing additional information on a topic encourages active thinking, which is an essential factor of fake news detection. However, descriptive knowledge is missing about users' behavior in social media with additional information at hand. Another issue regarding countermeasures such as source ratings is the dependence on active user engagement. Social psychologists suggest that the presence of social norms (SN) encourages people to behave in an intended way (Cialdini et al. 1990; Cialdini 2003; Goldstein et al. 2008). Concerning fake news in social media, however, it remains unclear whether SNs are an appropriate countermeasure in such digital environments. The doctoral thesis conducts three online experiments to investigate users' behavior in social media and draw conclusions from a sociocentric perspective. The results add to the body of knowledge in the fight against fake news and, therefore, also contribute to achieving sustainable development while avoiding intentionally misleading and, thus, counterproductive information.

Third, this thesis takes a technocentric perspective. In the course of this, technology is applied to achieve, for instance, structural and decision-making changes in organizations (Sarker et al. 2019; Pinsonneault and Kraemer 2002). Business enterprises, as a dominant form of social organizations, contribute to the damage and burden of our environmental nature (Melville 2010). IS has been one of the main reasons for increasing productivity in recent decades (Gholami et al. 2016). However, IS also have the potential to enable sustainable transformation in companies and to achieve environmentally sustainable businesses (Gholami et al. 2016; Melville 2010). From a technocentric perspective, suitable IS artifacts allow more efficient value creation processes to avoid waste. Waste, as a result of inefficiency, occurs not only when there is unnecessary overuse of resources, overstocking, or generation of garbage due to unused resources.

Waste also occurs when there are unexploited resources of time or inaccurate or unavailable information leading to false decisions. Therefore, IS artifacts can make use of the increasing penetration of digital technologies and the associated digitalization to avoid such waste.

Through the targeted use of digital technologies in IS artifacts, more economically and environmentally sustainable economies can be achieved. Evermore connected entities and the diffusion of digital technologies, innovative information flows enable various potentials in value creation processes and allow them to minimize waste. For instance, plentiful information availability allows production with fewer resources or the more accurate usage of perishable materials. However, evidence from the literature, as well as insights from research projects, illustrates that, to date, organizations struggle to apply digital technologies in a way to avoid waste. However, humans struggle to raise the opportunities of digitalization when developing value creation processes. One reason is that existing and established methods are not sufficient. This thesis outlines the need for advanced methods and presents a new approach to analyze and design lean, digitally supported value creation processes in companies and organizations. The presented Value Stream Modeling and Notation (VSMN) constitutes a domain-specific modeling language (DSML), which supports modelers and model users to design lean value creation processes concerning environmental and economic sustainability.

One form of social organization that is highly complex and, therefore, prone to waste are hospitals. Concerning rising costs for healthcare and the fact that hospitals account for approximately one-third of healthcare expenditures, especially in OECD (Organization for Economic Co-operation and Development) countries (Volland et al. 2017), hospitals have great potential for economic sustainability, which also has an impact on social sustainability (e.g., affordable healthcare). A significant expense driver in a hospital are logistics (Nachtmann and Pohl 2009) and, therein, especially the supply of materials (Ross and Jayaraman 2009). This thesis provides insights from research projects, which show that material logistic processes in hospitals suffer from insufficient or unavailable information and offer the potential to decrease waste. On these grounds, this doctoral thesis evaluates the utility, quality, and efficacy of VSMN using two examples. First, the thesis presents a reference model for material logistic processes in hospitals from a research and development project with two hospitals using VSMN for the design. Second, a material logistic process of another hospital is analyzed, and an appropriate target state is described using VSMN. The resulting artifacts illustrate the utility, quality, and efficacy of VSMN but also contribute to the body of knowledge for IS artifacts for hospital logistics. The knowledge fosters sustainable development regarding economic (i.e., fewer healthcare costs) and social (i.e., affordable healthcare) but also environmental dimensions (i.e., less resource consumption) from a technocentric perspective.

To summarize, the theoretical descriptive and prescriptive findings of this doctoral thesis add to existing design knowledge in the field of IS research for sustainability. The dissertation provides novel insights across the social-technical continuum to contribute to sustainable development. Therefore, the theoretical background on design knowledge and design artifacts in IS research is introduced throughout Chapter 2, followed by an introduction into methodological approaches within IS design science research. Chapter 3 takes a socio-technical perspective and provides design knowledge for EDFIS. Chapter 4 takes a sociocentric perspective and provides design knowledge about users' behavior in social media to counteract the societal threat of intentionally misleading false information. Chapter 5 takes a technocentric perspective and provides design knowledge to achieve lean, digitally supported value creation processes, among others, in hospitals. Finally, Chapter 6 draws on meta-inferences across these perspectives, discusses the results in the light of limitations, and provides an outlook for future research. Figure 2 summarizes the structure of this thesis and arranges Chapters 3 to 5 along the social-technical continuum.

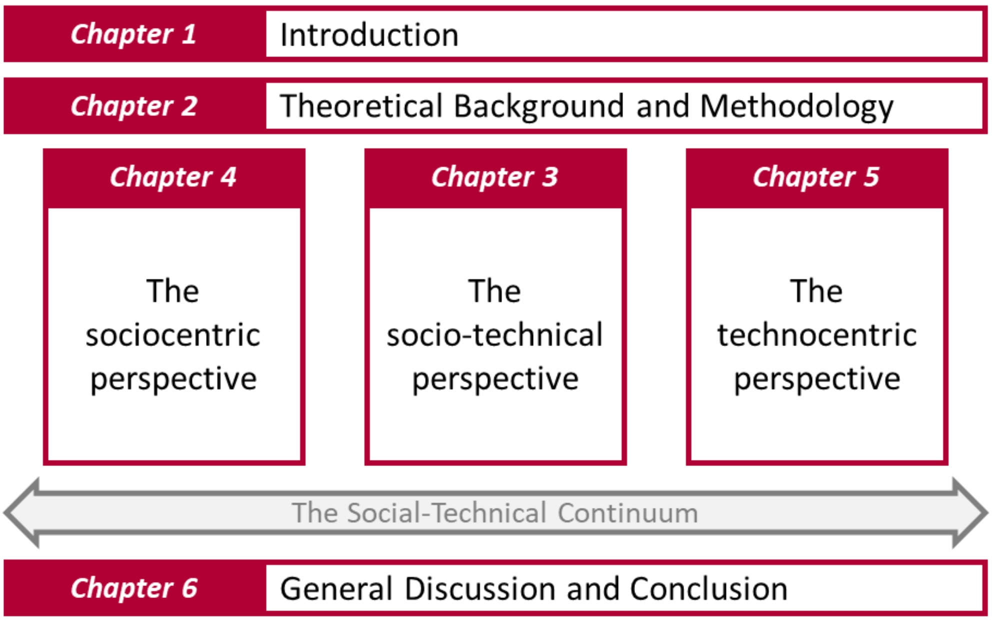


Figure 2: Structure of this doctoral thesis

During the research underlying this thesis, several parts of Chapters 3 to 5 were published in conference papers as part of a regular scholarly discourse or are under consideration for publication jointly with coauthors.[1] Major parts of Chapter 3 conform with Gimpel et al. (2020c) and Bätz et al. (2020). Major parts of Chapter 4 conform with Gimpel et al. (2020a) and Gimpel et al. (2020b). Major parts of Chapter 5 conform with Heger et al. (2020a), Heger et al. (2020b), Denner et al. (forthcoming), Heger et al. (forthcoming), and Thim and Heger (forthcoming).

[1] This doctoral thesis follows the "Promotionsordnung der Mathematisch-Naturwissenschaftlich-Technischen Fakultät der Universität Augsburg(in der Fassung vom 21.5.2014)" and the „Handreichung des Instituts für Materials Resource Management (MRM)für Doktorandinnen und Doktoranden zur Einbindung von Vorveröffentlichungen in eine monografische Dissertation im Rahmen einer Promotion zum Dr.-Ing. an der Mathematisch-Naturwissenschaftlich-Technischen Fakultät (MNTF) der Universität Augsburg (in der Fassung vom 09.01.2020)".

2. Theoretical Background and Methodology

2.1. Design Knowledge and Design Artifacts in Information Systems Research

IS research publishes knowledge involving two scientific paradigms: behavioral science and design science (March and Smith 1995). The behavioral science paradigm seeks to develop and justify theories to explain and predict organizational and human phenomena while analyzing, designing, implementing, and applying IS (Hevner et al. 2004). In contrast, the design science paradigm aims at deriving solutions that define the basis (i.e., practices, technical capabilities, and products) through which the lifecycle of IS (i.e., analysis, design, implementation, management, and use) can be effectively and efficiently accomplished (Denning 1997; Hevner et al. 2004). Design science, therefore, relies on existing natural or behavioral theories to derive artifacts (Hevner et al. 2004). Following Hevner et al. (2004), the result of design science is IT artifacts, which provide tools to extend the boundaries of human problem solving and organizational capabilities. Artifacts are something human-created instead of something which is naturally occurring (Hevner et al. 2004). Hevner et al. define IT artifacts "as constructs (vocabulary and symbols), models (abstractions and representations), methods (algorithms and practices), and instantiations (implemented and prototype systems)" (Hevner et al. 2004, p. 77).

A decade later, Lee et al. (2015) invoke to investigate increasingly IS artifacts instead of IT artifacts with the purpose of returning to the primal objective of IS being the design of IS. The authors, therefore, define IS artifacts as "a system, itself consisting of subsystems that are (1) a technology artifact, (2) an information artifact and (3) a social artifact, where the whole (the IS artifact) is greater than the sum of its parts" (Lee et al. 2015, p. 6).

First, a technology artifact is a tool, which is created by humans to solve human problems, achieve human goals, or serve human purposes. IT artifacts are, in part, a specialization of the generic technology artifact, which can be digital or electronic but also non-digital or non-electronic. A technology artifact must not necessarily be about information, nor being physical at all. Second, an information artifact is an instantiation of information, which occurs either directly or indirectly through human actions. Direct information is created, for instance, through human-spoken or written words. Indirect information instead is created through technology, for instance, a software application, which is executed by humans. The primal IT artifact, provided by Hevner et al. (2004), is thus divided into the technology and information artifact. Third, a social artifact includes not only individuals but rather relationships and interactions between or among humans to solve human problems, achieve human goals, or serve human purposes. Accordingly, the social artifact can include existing relationships as well as temporary interactions (Lee et al. 2015).

Design knowledge – with artifacts being a central component – results in design theory, as Gregor and Jones (2007) argue based on the seminal work of Simon (1996). Following Gregor and Hevner, “a theory is more abstract, has a nonmaterial existence, and contains knowledge additional to the description of a materially existing artifact” (Gregor and Hevner 2013, p. 341). Overall, five interrelated types of theory[2] are common in IS research to allow analysis, explanation, prediction, and prescription (Gregor 2006). Theory of design and action, as a theory which is informed by all other types of theory, “gives explicit prescriptions (e.g., methods, techniques, principles of form and function) for constructing an artifact” (Gregor 2006, p. 620). Design theories aim at proposing methodologies (for implementation of an artifact) or products (principles of form) in an abstract way, which could become physical with its instantiation. Constructs, models, and methods are abstract artifacts and, hence, components of design theory, while instantiations being material artifacts constitute a physical instance of theory (Gregor and Jones 2007). Altogether, a design theory consists of eight components: *purpose and scope*, *constructs*, *principles of form and function*, *artifact mutability*, *testable propositions*, *justificatory knowledge*, *principles of implementation*, and *expository instantiation*. While the first six components belong to the core of design theory, the latter two components are rather additional (Gregor and Jones 2007).

First, the *purpose and scope* define a set of meta-requirements to describe the system’s objective, scope, and boundaries. Those meta-requirements aim at developing a suitable design theory for a class of artifacts. While defining the objective of design theory, researchers rule out other objectives and, therefore, clarify the scope. Second, *constructs* are the most fundamental component of a design theory. They represent the entities of design theory, which are either physical or abstract and must be clearly defined. Constructs itself can state a more detailed design theory to handle complexity. Third, the *principles of form and function* describe the relationships between constructs and “define the structure, organization, and functioning of the design product or design method” (Gregor and Jones 2007, p. 325). These principles can occur,

[2] theory for analyzing, theory for explaining, theory for predicting, theory for explaining and predicting, theory of design and action

for instance, as an architectural blueprint but may also illustrate the sequence of a process. Fourth, *artifact mutability* addresses the need to consider in which ways artifacts can appear and develop over time. That is, artifacts appear to be in a current state of change. Besides, it is important to ensure that derived artifacts are also valid in different socio-economic contexts and practices and, thus, somewhat independent from a specific context. Fifth, *testable propositions* are some kind of measurable requirements. On the one hand, the propositions must allow ensuring that a derived design product meets its purpose and scope (i.e., it's meta-requirements). On the other hand, the propositions must allow ensuring that a design method is suitable for the implementation of an artifact according to the principles of form and function. Sixth, *justificatory knowledge* refers to theories that inform both, the design product and the design process, and constitutes the interconnection over all other components of a design theory. Such theories may be adapted from natural or social science or constitute other design theories. Justificatory knowledge must not be complete but available for the important parts of an artifact, which are mandatory for abstraction. Following Gregor and Jones' (2007) argumentation, it is inevitable to understand why an artifact works instead of proving that it is actually working. Seventh, the *principles of implementation* describe guidelines for the implementation of instantiations of design theory into our physical world. This, however, applies not only for the design product but also for the design process. Last but not least, Gregor and Jones (2007) argue to include an *expository instantiation* (i.e., prototypical instantiation) to provide a tangible example of something abstract like a design theory.

On these grounds, Gregor and Hevner (2013) argue that valid contributions from design science research projects can occur on different levels of maturity. These levels range from less abstract to more abstract artifacts. Less abstract contributions are instantiations of design products and processes (level 1). In contrast, nascent design theories (level 2), such as constructs, design principles, or models, built a step towards well-developed contributions in the form of design theories (level 3). Besides this, contributions in design science not only vary about the degree of abstraction but also regarding the maturity of the respective body of knowledge, which builds the foundation of design artifacts. Design science research utilizes and develops knowledge from two distinct knowledge bases.

On the one hand, descriptive knowledge describes natural, artificial, and human-related phenomena. For instance, such descriptions consist of observations, measurements, and cataloging into laws, principles, or theories. These phenomena and relationships offer the scientific base to understand our world. A frequently used term in design science is *kernel theories*, which describe “theories from natural or social sciences” (Walls et al. 1992, p. 43). On the other hand, prescriptive knowledge describes human-designed artifacts, which aim at enhancing our world. Following March and Smith (1995), constructs, models, methods, and instantiations constitute some kind of prescriptive knowledge. As design theories (Gregor 2006; Gregor and Jones 2007) include these types, they constitute the abstract and coherent body of prescriptive knowledge (Gregor and Hevner 2013). Contributions to the body of prescriptive knowledge depend on the maturity of the existing descriptive knowledge (application domain, describing how good the problem context is understood already) and prescriptive knowledge (solution domain, how well existing solutions already address the problem) (Gregor and Hevner 2013). Gregor and Hevner (2013) name three types of design knowledge contributions: *inventions*, *improvements*, and *ex-aptations*.

First, *inventions* are quite rare. Inventions constitute design products and design processes in some way, which has not yet been observed. The research activities, which result in inventions, must rather be exploratory and constitute a search for solutions in complex problem domains. Such activities do not follow a well-defined research process and require exceptional skills and knowledge of the researcher. Gregor and Hevner (2013) argue that invention artifacts can contribute to descriptive and prescriptive knowledge. Mostly, early contributions belong to level 1 contributions appearing as instantiations of radically new approaches. Consequently, newly invented artifacts add knowledge primarily to the body of prescriptive knowledge. Researchers are able to investigate the artifact when it is in use and derive new phenomena and sense-making relationships, which subsequently add to the body of descriptive knowledge. For inventions, no descriptive knowledge must exist yet. Otherwise, it is not an invention. However, descriptive knowledge from other disciplines may still inform the artifact design (Gregor and Hevner 2013).

Second, *improvements* are the most common contributions to design knowledge. Improvements constitute better-developed solutions in the sense of more efficient and effective artifacts. Such contributions occur when the solution domain’s maturity is on a low level, but the application domain is quite well explored. That is improvements built upon a large body of descriptive knowledge, while solution artifacts do not exist or are manifestly unsuitable. Hence, researchers must first describe why a new solution is necessary and built their argumentations on existing theories. Subsequently, researchers must evaluate new design artifacts to rigorously demonstrate improvements to achieve the design's objective (e.g., increasing efficiency). Improvements add to the body of prescriptive knowledge at all levels. Instantiations allow researchers to investigate improvements concerning existing instantiations of earlier artifacts. More clearly, nascent design theories (e.g., methods or models) add to the prescriptive knowledge base and also to the body of well-developed design theories. Besides, improvements may expand the body of descriptive knowledge as evaluations of improvements enhance the understanding of natural or social theories (Gregor and Hevner 2013).

Third, *exaptations* are the transfer of existing solutions to new problems. In this case, the understanding of a problem domain is rather poorly developed, and no (sufficient) solutions in the domain exist. Nevertheless, sometimes existing solutions from other application domains can be an effective solution and, therefore, can be transferred to the considered application domain. Often, newly invented technologies allow exaptations for existing problems. Hence, researchers expand the body of prescriptive knowledge while they prove that an exaptation is valuable and noteworthy. Contributions may occur on all three levels, ranging from instantiations to well-developed design theories. Moreover, exaptations may also add to the descriptive knowledge base as such contributions can provide a better understanding of the adapted artifact when it is applied to a new application domain (Gregor and Hevner 2013).

Beyond that, Gregor and Hevner (2013) explain that, theoretically, *routine designs* are the fourth type of contributions in design science. Routine designs constitute the application of known solutions to known problems when the solution domain and application domain draw from a sophisticated knowledge base. Thus, such contributions typically do not enhance the bodies of knowledge and are therefore left out.

Overall, design knowledge arises from the interplay of behavioral science and design science. Design artifacts and more mature design theories draw from so-called kernel theories but also help to understand natural and social phenomena better. As a scientific discipline at the intersection of behavioral sciences and design science, IS research makes essential contributions to derive solutions for existing problems but also contributes to a better understanding of our world we live in.

2.2.Methods in Information Systems Design Science Research

Behavioral and design science constitute complementary research cycles rather than two separate research paradigms (Hevner et al. 2004). For this reason, design science research aims at contributing to the body of prescriptive knowledge while also informing the body of descriptive knowledge. Researchers do have to draw from both knowledge bases to develop valuable new artifacts or investigate artifacts being in use (Gregor and Hevner 2013). Studies in behavioral science often implement artifacts to derive theories concerning the artifact’s use. For instance, field studies enable behavioral science researchers to better understand organizational phenomena in environmental settings. In contrast, design science aims at creating and evaluating artifacts as a possible solution for identified problems. Constructing and exercising artifacts allows design science researchers to understand the problem at hand and the feasibility of their artifact to its solution (Hevner et al. 2004).

With their seminal work, Hevner et al. (2004) provided guidelines on how to carry out design science research. Sound artifacts must illustrate the result of design science research projects. Those technology-based artifacts solve relevant problems and must be rigorously evaluated to demonstrate their utility, quality, efficiency, or other quality attributes. The evaluation of a design artifact – comparable to the justification of a kernel theory – requires well-defined and suitable evaluation metrics, which can be demonstrated using data from appropriate evaluation methods. For instance, observational methods such as case studies, analytical methods such as dynamic analysis, experimental methods such as studies in controlled environments, testing methods such as functional testing, or descriptive methods such as informed arguments are appropriate. Subsequently, the results of a research project must constitute unique contributions, which are verifiable. Therefore, rigor methods require to build contributions on the existing body of knowledge, which is comprised of kernel theories but also design knowledge. Design artifacts arise from an iterative search process, which is a cycle of designing and testing an artifact to discover suitable solutions. This development cycle is executed until a useful design artifact has been achieved. The problem at hand is that it is not clear what good means. One possibility is to calculate an optimal condition and to get as close as possible to it. Another possibility would be to compare an artifact's performance with existing solutions, which can be exceeded in terms of quality attributes regarding the design objective.

Building on these guidelines, Peffers et al. (2007) derive a research process for conducting and successfully presenting design science research. Their design science research method consists of six activities, which allow developing efficient design artifacts through an iterative search process following Hevner et al. (2004). These activities are *problem identification and motivation*, *define design objective*, *design and development*, *demonstration*, *evaluation*, and *communication*.

Overall, Peffers et al. (2007) structure their process sequentially but emphasize that iterations are possible. Also, it is not necessary to execute design science research projects in a strict order. The process by Peffers et al. (2007) starts with the step *identify the problem* within the application domain to develop an appropriate solution successfully. Therefore, it is advisable to grasp the problem conceptually and to elaborate on the added value of a solution. The latter is helpful to motivate the research project for the research community and to convey the own understanding of the problem. Second, researchers have to *define* either quantitative or qualitative design objectives. The definition of the objectives relies on the previous problem identification and knowledge about possible and suitable approaches to solve the problem. The objectives provide the foundation for the subsequent core of design science research. Third, the development of the artifact takes place throughout the *design and development* activities. Researchers specify the functionalities and the architecture of the design artifact at hand. In doing so, a researcher must involve descriptive knowledge (i.e., kernel theories) which appears suitable for a striven solution, but also consider prescriptive knowledge which describes existing artifacts. Subsequently, the fourth and fifth activities are closely related to each other. The *demonstration* requires appropriate knowledge of how to deploy the artifact to the physical world. Further, it allows a conclusion on the usefulness of the artifact. More significant, the *evaluation* examines the solution concerning its requirements and the intended design objective. Evaluation methods can both be qualitative as well as quantitative (as also depicted by Hevner et al. (2004)). Therefore, a researcher can adopt any evaluation approach to illustrate empirical evidence or logical proof (Peffers et al. 2007).

Peffers et al. (2007) do not provide detailed evaluation patterns or strategies to evaluate design artifacts rigorously. Yet evaluation activities in design science are important to add to both prescriptive and descriptive knowledge bases (Gregor and Hevner 2013). The missing or generic evaluation activity is specified, for instance, by Sonnenberg and Vom Brocke (2012) or Venable et al. (2016), who both are building on Pries-Heje et al. (2008).

Sonnenberg and Vom Brocke (2012) provide evaluation patterns that allow not only ex-post evaluation but also ex-ante evaluation. That is, artifacts should be evaluated before their actual instantiation and also, after they have been designed and demonstrated. In doing so, Sonnenberg and Vom Brocke (2012) leave behind the strict distinction of design and develop and, subsequently, demonstrate and evaluate as postulated by Peffers et al. (2007). Their evaluation patterns, moreover, are embedded at four distinct positions in the design science research process. First, after the problem identification, activities such as literature reviews, expert interviews, or surveys allow ensuring that the selected research problem at hand is meaningful and worth paying attention to. Second, after the design and development, activities such as assertions (i.e., making informed arguments), ex-ante demonstration (e.g., test case or analytical example), or focus group discussion allows ensuring that the designed artifact serves to solve the intended problem. Third, after the instantiation, activities such as demonstration with prototypes, experiments, or benchmarking aim at demonstrating the utility of an artifact at hand. Finally, after the application within the application domain, activities such as case studies, field experiments, or surveys allow to conclusively show the applicability and usefulness of an artifact in practice. The first three evaluation activities allow applying artificial methods (i.e., evaluating "a solution technology in a contrived and nonrealistic way" (Pries-Heje et al. 2008, p. 4)). Instead, the fourth evaluation activity only involves naturalistic methods (i.e., "explores the performance of a solution technology in its real environment" (Pries-Heje et al. 2008, p. 4)).

In the same vein, Venable et al. (2016) propose evaluation strategies to guide researchers while evaluating their results of the design and development activities. Therefore, the authors build on the same distinction into artificial and naturalistic evaluation paradigms as a dimension, which answers how to evaluate artifacts (Pries-Heje et al. 2008). Besides, the authors introduce a second dimension being the continuum between formative and summative evaluation. While formative evaluations ensure appropriate results of design research, summative evaluations compare the results with the intended objectives. However, most evaluation activities contribute to both purposes in some way. Venable et al. (2016) introduce four different evaluation strategies along both dimensions. First, a quick & simple strategy provides only a few formative evaluations. The strategy quickly moves to summative and naturalistic evaluation activities. Therefore, the approach is less effort and allows for fast results. Second, the *human risk & effectiveness* strategy utilizes formative evaluation methods early in the research process but moves quickly from artificial to more naturalistic approaches. The step from formative to more summative evaluation activities should be taken subsequently to include human and social challenges early in the research process. Third, the technical risk & efficacy strategy, vice versa, builds more intensively on formative evaluation methods. It takes the step to summative evaluation activities before also including naturalistic approaches. In this case, the benefit comes from the technical artifact more than other artifacts. Last but not least, the *purely technical* strategy should be applied for mainly artificial approaches as the intended point in time to deploy the artifact into the physical world has not yet been reached. Venable et al. (2016) also provide a four-step approach to choose a suitable evaluation strategy for a specific design science research project. Researchers first should define their evaluation goals and subsequently choose an appropriate strategy. Next, researchers should determine which properties have to be evaluated and, finally, design the distinct activities of their evaluation.

Getting back to the generic design science research method, introduced by Peffers et al. (2007), the authors finally propose to communicate the results of each design science research project. Following their suggestions, the results of design science research should not only be addressed to researchers but also to practitioners. More precisely, in IS research, the audience of design knowledge contributions consists of both, technical-oriented and managerial-oriented audiences.

Overall, there are different starting points to initiate design science research projects. The problem-centered approach starts with the first activity in the research process and appears when researchers, for instance, observe a problem in the real world. The objective-centered approach arises from a specific need to employ a suitable artifact. Therefore, the research process would start with the second activity. The design and development-centered approach, beginning with the third activity, is proper when an appropriate artifact already exists but has not yet been investigated as a solution for a given problem. Finally, the client/context-initiated approach starts with the discovery of an appropriate solution, which has been deployed in practice already. In this case, researchers would work backward and develop the abstract design ex-post to its instantiation (Peffers et al. 2007).

The doctoral thesis at hand builds on these methods for design science research and aims at contributing prescriptive as well as descriptive design knowledge along a social-technical continuum. By this, the thesis' overall objective is to add to the body of knowledge regarding the suitable design of IS to enable sustainable development. Therefore, the thesis applies qualitative as well as quantitative methods. Following the seminal work of Venkatesh et al. (2013), multi-methodological perspectives allow deriving rich insights into different types of phenomena. For instance, qualitative as well as quantitative methods in combination allow complementary insights into solution and application domains as well as deriving a clearer understanding of a phenomenon under consideration. Mixed methods are also applied to expand the body of knowledge from a different point of view and allow researchers to gain more divergent views of a phenomenon.

Chapter 3, mainly builds on quantitative methods to gain prescriptive and descriptive knowledge from a socio-technical perspective. Similarly, Chapter 4 derives descriptive knowledge from a sociocentric perspective applying primarily quantitative methods. Chapter 5 utilizes mostly qualitative methods to gain prescriptive knowledge from a technocentric perspective.

3. Design Knowledge from a Socio-Technical Perspective

Green IS are a means to enable environmentally friendly behavior and decision-making on an individual level (Henkel and Kranz 2018; Watson et al. 2010). To address the individual level, IS scholars call to develop knowledge to improve action formation (Henkel and Kranz 2018). Improving individual behavior by means of the targeted application of green IS requires to develop solutions at the interplay of the technical and social systems (Sarker et al. 2019). Therefore, this chapter takes a socio-technical perspective to improve pro-environmental behavior. A stream within green IS relates to the application of IT to increase energy efficiency in the sense of decreasing CO_2 emissions (Watson et al. 2010). The transportation sector, as the third-largest emitter, accounts for 23% of global GHG emissions (IPCC 2014) and, as a part of the sector, individual traffic emits 11% of the world's total CO_2 emissions (Andor et al. 2020). The sector thus presents an important context for pro-environmental behavior on an individual level. Chapter 3.1 presents prescriptive knowledge for eco-feedback IS to improve driving behavior. Subsequently, Chapter 3.2 presents descriptive knowledge about improvements in driving behavior due to eco-feedback. Major parts of Chapter 3 conform with Gimpel et al. (2020c) and Bätz et al. (2020).

3.1. Designing Mobile Eco-Driving Feedback Information Systems – Results from a Field Study to Improve Driving Behavior

As already mentioned, traffic causes substantial GHG emissions. Besides technological improvements (e.g., more fuel-efficient engines), there is substantial unused potential for emission reduction through people's behavior. The choice of climate-friendly transportation, carpooling, purchase decisions for fuel-efficient or electric motorization, and driving behavior are starting points. According to Lárusdóttir and Ulfarsson (2014), the latter has a significant influence on the fuel consumption of vehicles per driven kilometer through acceleration, deceleration, and average speed. Hence, one way to achieve reductions in fuel consumption and, subsequently, in CO_2 emissions is through a change in the behavior of car drivers. From an IS perspective, human behavior can be improved by giving feedback (Froehlich et al. 2010). Based on additional available information, drivers can adapt to their behavior. As different studies show, it is possible to reduce fuel consumption through feedback systems by between 1% and 7% on average (e.g., Tulusan et al. 2012a; Boriboonsomsin et al. 2010).

Current research strands in IS research aim at contributing to sustainability, one of which is *IS for Environmental Sustainability* (Melville 2010; Watson et al. 2010). The strand investigates IS usage to achieve environmental practices, also known as green IS (Henkel et al. 2017; Henkel and Kranz 2018; Watson et al. 2010). Green IS do have the potential to foster pro-environmental behavior (Henkel and Kranz 2018). However, there is still a need for solution-oriented studies to leverage IS to achieve environmental sustainability (Gholami et al. 2016).

Over the last decades, much work has been done to foster pro-environmental behavior utilizing eco-feedback (Fischer 2008; Karlin et al. 2015). With smart, connected cars and more embedded sensors, EDFIS become possible to offer ever-richer but also highly individual feedback to drivers. Such EDFIS are promising tools to contribute to a more sustainable lifestyle. Several studies have examined behavioral improvements using eco-feedback to save fuel (e.g., Boriboonsomsin et al. 2010; Tulusan et al. 2012b, or Chapter 3.2 of this thesis). Mobile devices are becoming increasingly popular for EDFIS but also in the mobility sector (Kamilaris and Pitsillides 2016). Mobile devices like smartphones with their sensors, computing power, internet connectivity, user interface, and wide-spread availability are very well suited to give users individual feedback on their driving behavior (Froehlich et al. 2009). However, research has rarely considered mobile-only solutions for EDFIS. We have found only two corresponding contributions (Tulusan et al. 2012a; Dahlinger et al. 2018). Indeed, mobile EDFIS can make an essential contribution to the avoidance of CO_2 emissions, especially in emerging trends such as car sharing (Shaheen and Cohen 2013). But to date, no compilation of design knowledge exists that describes the structure and principles of form and function of such a system. Therefore, we aim at expanding the body of design knowledge for mobile EDFIS. We pose the design objective (DO) as follows:

DO 3-1: Design a mobile eco-driving feedback information system that supports car drivers individually in adopting a more environmentally sustainable driving behavior.

With this, we aim for design knowledge of a mobile EDFIS that satisfies two design requirements: firstly, improving fuel-efficient driving behavior to reduce fuel consumption over the same distances traveled and, secondly, developing an artifact being an improvement concerning the effectiveness of existing EDFIS artifacts (Gregor and Hevner 2013). Regarding the latter, our reference is the mobile EDFIS artifact, which has been investigated by Tulusan et al. (2012b). We do not address the broader scope of influencing, for instance, the choice of a mode of transportation (Froehlich et al. 2009). Neither do we consider technological solutions that may reduce fuel consumption (like autonomous vehicles). With the advent of self-driving cars, electric vehicles, and other technological innovations, individual driving behavior will likely lose importance in reducing fuel consumption and CO_2 emissions. Although these approaches, compared to behavioral changes, can have a higher leverage effect on CO_2 emissions, we focus on improving driving behavior, which can quickly show initial successes without renewing the pool of cars being on our streets today and likely for several decades to come. We posit that the behavior of individual car drivers will remain relevant for quite some time, as will the search for a means to influence this behavior positively.

This chapter fills the described gap, applying the design science research methodology (Peffers et al. 2007; Hevner et al. 2004) and is structured as follows: First, we introduce our design science research process. Next, we summarize the theoretical background and related academic work. On these grounds, we develop design principles and propose an architectural blueprint. Subsequently, we demonstrate and evaluate our findings using a prototypical instantiation within a field study. Finally, we conclude with a discussion of implications and limitations.

Methodology

Today, much knowledge from behavioral science exists about the effect of feedback on improving eco-driving. On these grounds, we apply the design science research methodology, as proposed by Hevner et al. (2004) and Peffers et al. (2007). Figure 3 summarizes the emergence of our design knowledge.

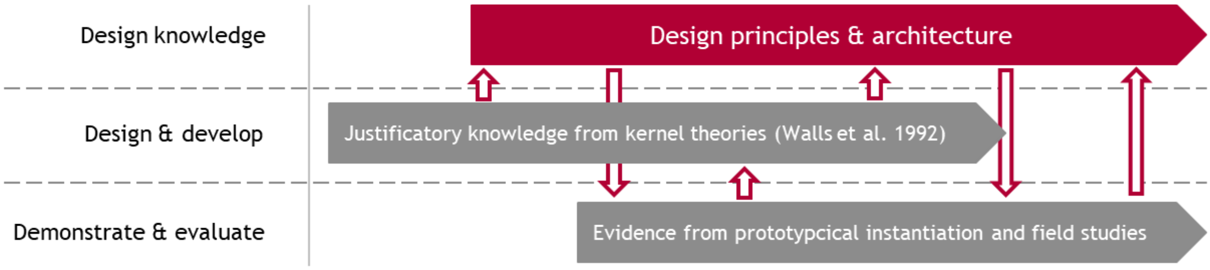


Figure 3: The design knowledge develops throughout an iterative search process

We already defined our DO 3-1 and corresponding design requirements. Subsequently, we introduce existing justificatory knowledge on driving behavior, eco-feedback, and existing EDFIS. On these grounds, we derive theoretical design knowledge – consisting of the design principles (DP) and an architectural proposal of the artifacts – from an iterative search process (Hevner et al. 2004). Following Hevner et al. (2004, p. 88), “design science is inherently iterative. The search for the best, or optimal, design is often intractable for realistic IS problems.

Heuristic search strategies produce feasible, good designs that can be implemented in the business environment." We executed two research cycles and used insights from the subsequent demonstration and evaluation phase to improve our design artifact, as suggested by Hevner et al. (2004). We communicate mature design knowledge (after 2nd iteration) for mobile EDFIS with this chapter, which has been successfully demonstrated and evaluated using a prototypical instantiation within a field study. The first iteration is reported in Appendix A. The instantiation implements the presented design principles and proposed architecture. Data gained from the field study suggests that our design knowledge describes an effective solution, which achieves DO 3-1 and its corresponding design requirements.

Theoretical Background

Eco-Driving

Driving behavior is a complex behavior, which is depicted in two fundamental aspects. Firstly, strategic driving behavior depicts overarching decisions. It is also referred to as travel behavior. It includes, among other things, the chosen route and trip goals such as minimizing time or costs but also the choice of transportation mode (Michon 1985; van der Molen and Botticher 1987). The second aspect depicts execution-related driving behavior. It refers to tactical and operational driving behavior, including driver attitudes such as calm or aggressive driving behavior (Michon 1985; Rafael et al. 2006). Studies have found that calm driving behavior is characterized by a low gear-shifting frequency, slow acceleration, and driving speeds not exceeding the legal limit. Aggressive driving, on the other hand, involves a higher tendency to shift gears, hard acceleration, and speeds above the legal speed limit (Rafael et al. 2006).

Prior research also found evidence that fundamental aspects of driving behavior have a significant effect on fuel consumption and, thus, lower CO_2 emissions. Ericsson (2001) found 62 driving parameters, which can be aggregated to 16 independent factors describing operational driving behavior. Of these, moderate and hard acceleration, a strong speed oscillation, many stops during a trip, and late gear changes from gear 2 to 3 increase fuel consumption. On the other side, deceleration, driving speed between 50 and 90 km/h, moderate engine speed at gears 2 and 3, as well as low engine speed at gears 4 and 5 decrease fuel consumption. These results are also consistent with the work of Lárusdóttir and Ulfarsson (2014). Their findings provide evidence that the driven distance, hard acceleration events, and somewhat higher average speed increase fuel consumption of vehicles per kilometer driven. Similarly, the results provide evidence that higher fuel consumption also results from hard deceleration events, numbers of stops, and idle time during trips (Lárusdóttir and Ulfarsson 2014). On these grounds, eco-driving behavior is an appropriate lever to reduce fuel consumption and thus contribute to the fight against rising CO_2 emissions.

Eco-Feedback and Prior Research on Fuel Consumption

A prominent approach to address behavioral improvements is feedback. Feedback is a "communications process in which some sender (…) conveys a message to a recipient (…) [that] comprises information about the recipient" (Ilgen et al. 1979, p. 350). According to feedback intervention theory (Kluger and DeNisi 1996), this information enables the creation of a gap between a person's behavior and some standard or individual goal, resulting in a person's desire to reduce this gap. In the context of pro-environmental behavior, this is often referred to as eco-feedback, which is defined as "feedback on individual or group behaviors to reduce environmental impact" (Froehlich et al. 2010, p. 1999). The effectiveness of the eco-feedback heavily depends on what information is displayed and how (Froehlich et al. 2010).

Researchers have employed a range of perspectives to examine the impact of eco-feedback on fuel consumption over the last decades. Those studies have investigated the effect of eco-feedback using different EDFIS. On the one hand, eco-feedback has been applied to improve rather strategic driving behavior, such as reducing car usage and, therefore, annual mileage at all. For instance, Graham et al. (2011) find a positive effect when providing eco-feedback on environmental and financial savings (CO_2 and money) to a group of students while they do not use their cars. On the other hand, various studies investigated improvements in fuel consumption due to rather execution-related driving behavior, which is most important when it comes to fuel consumption per driven distance (see our first design requirement).

First, in 1989, feedback with other information, task assignment, and control were considered as influencing factors to reduce one's energy consumption (Siero et al. 1989). The study has investigated drivers in a business context. The EDFIS was non-digital as they provided eco-feedback employing a bulletin board. Their results achieve fuel savings of up to 7.3% (Siero et al. 1989). Siero et al. (1989) also applied SNs to the drivers besides eco-feedback, which allowed comparisons among the drivers. van der Voort et al. (2001), instead, conducted an experiment using a driving simulator. The system provided the subjects with their fuel consumption based on their actual driving behavior. Their results similarly indicate fuel savings of up to 7% (van der Voort et al. 2001).

With the spread of digital technologies, EDFIS also have changed. The application of digital technologies allows investigating the effect of eco-feedback rather directly and more precisely in naturalistic settings. Today, several cars' on-board systems contain recommendations like the most fuel-efficient gear (Metzler et al. 2013). The introduction of the so-called on-board diagnostic II interface (OBD-II) allowed collecting data from the vehicle's sensors during a trip externally. Boriboonsomsin et al. (2010) investigated the effect of eco-feedback using an OBD-II enabled device. Their results from a study with 23 participants provide more detailed insights. Eco-feedback on actual fuel consumption and CO_2 emissions improved driving behavior to achieve savings between 1% (highway-context) and 6% (city-context) (Boriboonsomsin et al. 2010).

Tulusan et al. (2012b) conducted a field study using a smartphone application to present feedback on driving behavior to 50 corporate drivers. Their results show improvements in fuel efficiency by savings of 3% on average. Their mobile EDFIS provided eco-feedback about operational driving behavior, such as acceleration or speed (Tulusan et al. 2012b). Furthermore, the authors derived that car drivers prefer direct feedback on operational driving behavior during the trip instead of indirect feedback afterward (Tulusan et al. 2012a).

Kurani et al. (2013) achieve similar fuel savings in a real-world scenario using a recording and display device. Interestingly, a survey of the participants in the study showed that the drivers had little knowledge about efficient driving behavior. Few could name specific points they could improve to save fuel (Kurani et al. 2013). Therefore, Vagg et al. (2013), for instance, introduce a driver assistance system to 15 vehicles. Their instantiation provides information on inefficient acceleration and early upshifting the gears to achieve fuel-efficient driving behavior.

Similarly, Magana and Organero (2013) did not provide eco-feedback on the actual outcome of driving behavior. Instead, their study provided an eco-score, which was compared to other drivers and thus constituted a gamification element based on social comparison. Drivers successfully reduced their fuel consumption over time. Ecker et al. (2011) also find an indication that social comparison leads to increased energy efficiency for 37 drivers. Dahlinger et al. (2018) derive evidence from their results of 62 road assistance drivers that symbolic eco-feedback (i.e., a tree growing or withers) achieves fuel savings between 2-3%. They conclude that practitioners have to consider the design of the provided feedback carefully to implement effective EDFIS.

Overall, eco-feedback affects fuel consumption and fosters fuel savings through influencing driving behavior (insights to behavioral-specific improvements, see Chapter 3.2). However, changing driving behavior not necessarily achieves fuel savings (for instance, Ecker et al. (2011) find an impact on driving behavior but no fuel savings). Consequently, the design of mobile EDFIS must address driving behavior efficiently. To date, only Tulusan et al. (2012b) and Dahlinger et al. (2018) applied mobile solutions. Their instantiations (following the wording of Hevner et al. 2004) achieve fuel savings of approximately 3% on average.

Design of Mobile Eco-Driving Feedback Information System Artifacts

The design artifacts are design principles (describing the principles of form and function following Gregor and Jones (2007)) and a proposed architecture for mobile EDFIS. The design artifacts are based on justificatory knowledge from literature.

Design Principles

The following aspects are essential, according to Froehlich et al. (2010) and Paay et al. (2013), to design effective eco-feedback systems: information, feedback, mobility, expert advice, self-comparison, and community information (for the latter two also change over time).

First, *information* about the individual behavior of the recipient (i.e., the driver) to assess one's behavior must be collected to affect future decisions. It needs to be displayed in a way, which is easy to understand, attention-grabbing, trustworthy, memorable, and presented at the time of decision (Froehlich et al. 2010; Brewer and Stern 2005; Geller et al. 1982). When individuals are motivated to improve their behaviors, displaying information has an educational effect and raises awareness (Schultz 2014).

From the mere display of information, an individual does not necessarily recognize the gap between the own behavior and the desired behavior. It is essential to highlight this gap by providing *feedback* at the right moment (Allison and Stanton 2019). Feedback can vary from a high-level to a detailed one (Karlin et al. 2015; Froehlich et al. 2010) to draw a person's attention to a specific problem and, thus, encourages to consider how a person's behavior may contribute (Fischer 2008). Detailed feedback provides dedicated information on how to change a particular behavior. High-level feedback demonstrates the consequences of one's actions (Karlin et al. 2015). One example is environmental damage through high fuel consumption. Feedback can raise people's awareness of the relevance of their behavior. Likewise, it can increase people's understandings of the consequences of behavioral change but therefore need to be connected to the moment of the decision causing a particular behavior. In this way, feedback makes clear the links between the actions of individuals and the problem at hand, for example, by explaining the increases and decreases in energy consumption that result from specific behavior (Fischer 2008). Hence, as with information, feedback needs to be connected to the dedicated behavior (Fischer 2008; Froehlich et al. 2010). Altogether this can be summarized as our first design principle:

DP 3-1: Provide information about driving behavior and give eco-feedback on relevant aspects of driving behavior to highlight the gap between the driving behavior of an individual and the desired driving behavior.

Closely related, Fischer (2008) found that providing eco-feedback is more effective when individuals receive not just one type of eco-feedback but receive multiple types of feedback like consumption over time, environmental impact, or saving tips. Consequently, a combination of numerous ways to give feedback seems to be favorable (Froehlich et al. 2010).

DP 3-2: Provide multiple types of eco-feedback as this enables the individual to identify the most propelling information with regards to their driving behavior.

In the context of pro-environmental behavior, literature differentiates between two concepts why people adopt eco-friendly behavior (Froehlich et al. 2010). First, rational choice models assume that pro-environmental behavior is driven by self-interest, primarily by systematic processes evaluating expected utility. Specifically, the rational-economic model postulates that people act to maximize the benefits or minimize the expenses. As far as the environment is concerned, this model is likely to be simplified to suggest that people will adopt economically beneficial and environmentally responsible behaviors[3] (Froehlich et al. 2010). Different studies on the application of eco-feedback regarding the economic impact of behavior have already proven its effectiveness. Those studies mostly provide economic feedback on energy costs like fuel or electricity costs (e.g., Brandon and Lewis 1999; Boriboonsomsin et al. 2010) or monetary savings (Dalén and Krämer 2017). Graham et al. (2011) provide additional economic feedback on the financial savings associated with avoidance of car driving at all.

[3] Costs may not always be of financial nature. Regarding the environment, the emission of GHG can also be considered as costs.

Consequently, the economic impact (i.e., the costs) of driving behavior should also be mirrored to the driver by a mobile EDFIS in the context of eco-driving. Others, however, are less motivated about monetary savings. For instance, people with an environmentally friendly mindset may be encouraged by highlighting the environmental impact of their behavior (or vice versa, the environmental protection due to their behavioral improvements).

Existing studies provide eco-feedback about CO_2 emissions (Graham et al. 2011), or consumed energy and emitted CO_2 simultaneously (Brandon and Lewis 1999). In the context of eco-driving behavior, the most common EDFIS is the fuel consumption meter, which is effective in itself (Hiraoka et al. 2009; Hiraoka et al. 2010). Boriboonsomsin et al. (2010) provide feedback on fuel consumption but also emitted CO_2. Dahlinger et al. (2018) and Kurani et al. (2013) investigate feedback on fuel consumption concretely and abstractly. Consequently, addressing cost awareness as well as environmental awareness, are proven motivational aspects to be effective. In summary:

DP 3-3: Provide feedback on the impact of driving behavior, encompassing both the financial (i.e., economic) as well as the environmental impact of driving behavior to create awareness for the consequences of the individual's driving behavior.

Second, norm-activation models acknowledge that "personal norm activation […] may trump subjective perceptions of utility" (Staats et al. 2004; Froehlich et al. 2010, p. 2001). That is, people perhaps do not make behavioral choices that maximize personal utility but make behavioral choices as they are feeling SNs require them to do so. Research provides evidence that SNs have a significant influence on people's decision making (Cialdini et al. 1990; Deutsch and Gerard 1955). Such normative messages can either provide information about behavior, which is approved or disapproved, or information about the behavior of others, which is likely to be effective (Cialdini et al. 1990). As long as the rational choice for an individual is also the rational choice for the collective, the effect of eco-feedback is not altered if an individual has additional information about the comparable behavior of peers. The person is triggered in both cases the same way, and no conflict of interest arises. In the case that the individual rational choice is not in the interest of the collective, additional information about comparable behavior of peers will alter the individual's decision in favor of the collective (Froehlich et al. 2010). Consequently, another effective way to foster improvements towards desirable behavior is by providing information about peers in a similar situation (Froehlich et al. 2010; Paay et al. 2013; Schultz 2014). *Community information* potentially triggers individuals even more to adjust their behavior. The community serves as a comparison and may vary from anonymous consumers (Brandon and Lewis 1999; Goldstein et al. 2008; Magana and Organero 2013) as well as to known consumers like neighbors (Schultz et al. 2007), colleagues (Siero et al. 1996), or friends (Jain et al. 2012; Foster et al. 2010). A popular way to carry out social comparison are rankings (Ecker et al. 2011; Magana and Organero 2013). Altogether:

DP 3-4: Provide normative feedback on the driving behavior of a peer community to compare it to individual driving behavior, as SNs serve as additional motivation.

Besides motivation, providing eco-feedback over a long period is likely to ensure habit formation and a long-lasting effect. Darby (2006) and Fischer (2008) find that comparing current consumption to previous consumption is a useful instrument to decrease energy consumption. *Self-comparison* of own performance helps to assess personal behavior and promotes change. People are becoming aware of effective and efficient behavioral improvements. Furthermore, constant comparison against previous behavior promotes the long-term adoption of desirable behavior (Froehlich et al. 2010; Paay et al. 2013; Riche et al. 2010). In the context of eco-driving, constant self-comparison is increasing utility and acceptance of EDFIS. This holds particularly true for learning-oriented drivers (Brouwer et al. 2015). Thus:

DP 3-5: Provide information on the driver's performance over time to monitor the effectiveness of changes in driving behavior performed by the individual to increase learning.

Another aspect of adapting more environmentally friendly driving behavior is guidance. Evidence from existing studies illustrates that drivers do not necessarily know how to improve their driving behavior to achieve fuel savings (Kurani et al. 2013). As a consequence, drivers who receive advice regarding potential improvements were able to increase their fuel efficiency significantly throughout multiple studies (e.g., van der Voort et al. 2001; Tulusan et al. 2012b). Insights about eco-feedback to increase energy conservation supports these findings. People do need guidance on which behavioral aspects they should focus on to improve energy-efficient behavior (Karlin et al. 2015). A first step is providing information regarding the consequences of people's behavior, so they become aware of the fact that an improvement of their behavior is necessary. Merely being aware of the need to change, however, is not enough (Schultz 2014). Hence people must be guided in which way their behavior needs to change. A feasible way to do so is by giving *expert advice* to the individual as this provides sound knowledge and suggests a trustworthy option (Paay et al. 2013).

DP 3-6: Provide expert advice to enrich eco-feedback, as this outlines entry points for improvements in driving behavior and signals the individual that the information is trustworthy.

The medium providing eco-feedback to the user is most effective using interactive computerized devices as "interactivity, and the possibility of choice involve [recipients], raise their attention and allow for tailored solutions" (Fischer 2008, p. 99). Interactive feedback systems motivate users to deal with their behavior. Ueno et al. (2006) find interactive IS motivate to reduce energy consumption. Karlin et al. (2015) support this as their results showed the largest effect sizes when studies provided feedback using engaging or interactive media. An interactive feedback system may provide rich data that can be analyzed from different perspectives and with the help of various statistics. Still, people must achieve a sense of reflection on their behavior that influences consumption. People must be able to link their behavior to its consequences with the help of behavior-related feedback (Fischer (2008) speaks of *appliance-specific feedback* in the context of electricity consumption). As smartphones have become our daily companions, no other technical device is more closely connected to our lives than the smartphone.

Consequently, a feedback device needs to be *mobile* to show information at the right moment (i.e., in real-time) and highlight the gap between current and desired behavior (Paay et al. 2013; Tulusan et al. 2012a; Lange and Dewitte 2019). Especially in the context of carsharing or other forms of multiple people using the same car (e.g., a family car) or individuals using more than one car (e.g., multi-car households or a private car and a company car), EDFIS need to be truly mobile and must not be bound to a specific car or any hardware which is not available in every car. Henceforth, the smartphone's sensors pose a viable opportunity for giving personalized feedback regardless of the driven vehicle. This results in:

DP 3-7: Provide eco-feedback via a mobile device as this grants flexibility to the individual seeking for eco-feedback.

Abstract Blueprint

Seven design principles summarize the design knowledge for mobile EDFIS derived from existing literature. Subsequently, we derive a proposal for an architecture, which implements the design principles. The resulting *abstract blueprint* (Gregor and Jones 2007) illustrates the high-level architecture of a socio-technical IS artifact (Bostrom and Heinen 1977; Briggs et al. 2010;

Sarker et al. 2019), which describes the interplay of the technical system (being the feedback system) and the social system (being the drivers). Figure 4 summarizes the blueprint.

The technical system (Bostrom and Heinen 1977) includes three layers. Each design principle impacts one or more layers. First, the sensor layer allows capturing driving behavior of an individual driver, including behavioral aspects of driving, such as acceleration and speed. Besides, the layer provides functionality to collect data on the actual outcome of driving behavior, which is the fuel consumption of the car. Second, the feedback layer derives feedback information from the gained data.

On the one hand, behavior-related feedback allows the driver to recognize and adjust inefficient behavior while driving. On the other hand, rational and normative feedback addresses different motivational aspects and thus is likely to redirect the driver’s attention to the EDFIS. For the calculation of the normative feedback, the information about the driving behavior of other drivers is required. Therefore, each EDFIS has an interface to enable data exchange. Such data exchange can, for example, be implemented via a central platform (cloud). Third, the presentation layer provides feedback to the driver. Behavior-related feedback is provided to the driver in real-time. Furthermore, rational and normative feedback is provided to the driver after a trip. The EDFIS compares the actual feedback with recent trips and also allows long-term comparisons.

The social system includes individual skills, relationships among individuals, and rewards for individuals (Bostrom and Heinen 1977). Concerning mobile EDFIS, first and foremost, the individual driver, whose driving behavior is recorded and who receives personal feedback, can adapt to environmentally friendly driving behavior and, at the same time, receives motivational information (rational and normative) to improve driving skills. Second, the normative feedback addresses the relationship between drivers and allows social comparison. Third, rational feedback addresses financial (i.e., fewer fuel costs) and ecological (i.e., healthier environment) rewards. Together, these aspects constitute the social system, which interacts with the technical system.

To sum up, we derived seven design principles from literature, which describe the principles of form and function following Gregor and Jones (2007). On these grounds, we propose an architecture for mobile EDFIS. Subsequently, we demonstrate and evaluate the developed design knowledge to ensure its utility, quality, and efficacy (Hevner et al. 2004).

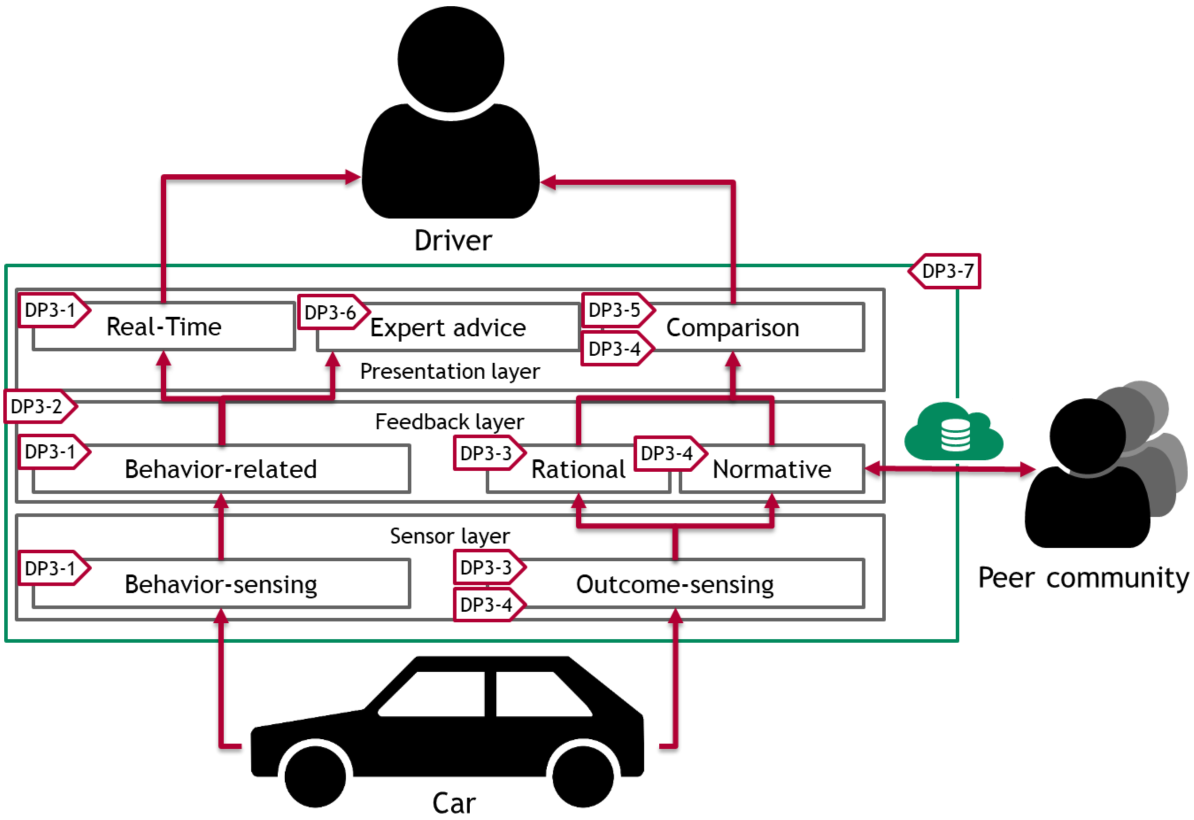


Figure 4: The architecture of the mobile EDFIS

Demonstration and Evaluation

"The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods" (Hevner et al. 2004, p. 83). We develop the evaluation of the presented design artifacts following Venable et al. (2016). The evaluation's goals are to demonstrate the utility, quality, and efficacy of the design artifacts, and the fulfillment of the design requirements (step 1 of Venable et al. (2016)). Our proposed design artifacts constitute a socio-technical IS artifact and thus require human interaction. For this reason, we choose a human risk and effectiveness evaluation strategy (step 2), since "the major design risk is social or user-oriented (…) [and] a critical goal of the evaluation is to rigorously establish that the utility or benefit will continue in real situations" (Venable et al. 2016, p. 6). The evaluation properties (step 3) are the design objective (DO 3-1) and the corresponding design requirements. Four episodes (step 4) constitute the basis of the evaluation: justification from the theoretical background (see above), the prototypical instantiation of a smartphone application, beta tests to gain feedback, and a field study as a major evaluation method (see following subchapters).

Prototypical Instantiation of the Mobile EDFIS Design

Initially, we implement an instantiation of the design artifacts to demonstrate the design principles and the proposed abstract blueprint.[4] The prototype originates as an Android smartphone app, with advantages that include Android's high market share, easy access to the app store, many open-source frameworks, and sensors associated with the smartphone (DP 3-7). Our prototype uses the device's GPS (global positioning system) sensor to track driving behavior (speed and distance). Further, we access the accelerometer sensor to discover the strength of the driver's maneuvers (DP 3-1).

In addition to the consumption meter, the app provides environmental feedback as emitted CO_2 in kilograms per 100 kilometers compared to the last trip, as well as the absolute number of emitted kilograms during this trip. At the same time, it provides economic feedback in terms of the fuel costs per 100 kilometers compared to the last trip and the total fuel costs during this trip. The driver is encouraged to enter the fuel consumption at the end of the trip, as displayed by the car's consumption meter.[5] The app then calculates the environmental and economic feedback (DP 3-3). Additionally, the app provides a social comparison that the app calculates from the consumption input. This feedback is presented as a ratio of drivers with similar engines but lower consumption than the driver who uses the app. Thus, the app provides information about how many other drivers do better regarding their fuel consumption (DP 3-4). The prototype uses a database containing frequency distributions fetched from the online platform *spritmonitor.de* to calculate this ratio. The app provides six distributions for both petrol and diesel engines: 50 horsepower (hp) or less; 51 to 100 hp; 101 to150 hp; 151 to 200 hp; 201 to 250 hp; and 251 hp or more. Drivers enter their car's fuel type and horsepower once during the initial usage of the app. On these grounds, the app compares the driver's fuel consumption to the consumption of the respective reference group.

The feedback is provided to the driver at different points in time. On the one hand, the driver receives a driving score as behavior-related feedback in real-time to adjust driving behavior while driving (DP 3-1). On the other hand, the application provides economic and environmental feedback, as well as a social comparison to the driver subsequently to the trip (DP 3-2 and

[4] There was an earlier version of the app, which has been developed throughout the iterative search process. For clarity and brevity, we do not elaborate on the intermediate design artifacts and present only the mature instantiation. The Appendix A provides details about the earlier version.
[5] This manual input could be obtained automatically from the car via the OBD-II interface. For demonstrating the design, we spared the hardware and software necessary for the interface.

DP 3-3). In addition, the app provides a summary of behavior-related driving behavior: mean speed and maximum speed, duration, and distance of the trip. All values are displayed in comparison to the previous trip. Improved values (in terms of being associated with less fuel consumption) are marked green, while worse values are marked red to reward better driving behavior (DP 3-5).

The app provides historical data in the form of charts that allow interaction and enable comparison of current and previous behavior over a long-term period (DP 3-5). Besides, the app offers a range of expert saving tips to help users improve their driving behavior (DP 3-6). Figure 5, Figure 6, and Figure 7 present screenshots of the app.

Prior to publication, we tested the app with seven beta testers. They helped us to improve the usability and intelligibility of the app but also offered first feedback on our design principles. The app has an upload functionality that sends all collected data to our file storage. Overall, our experience is that smartphone usage and rich options to design feedback (e.g., text, graphics) are suitable for the provision of individual feedback.

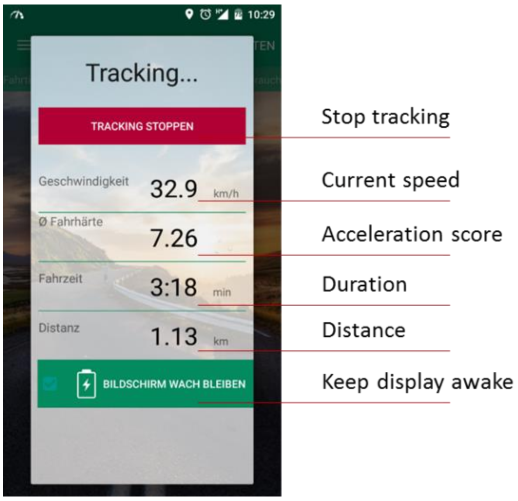


Figure 5: Screenshot of the abstract feedback during the trip (German original with translations)

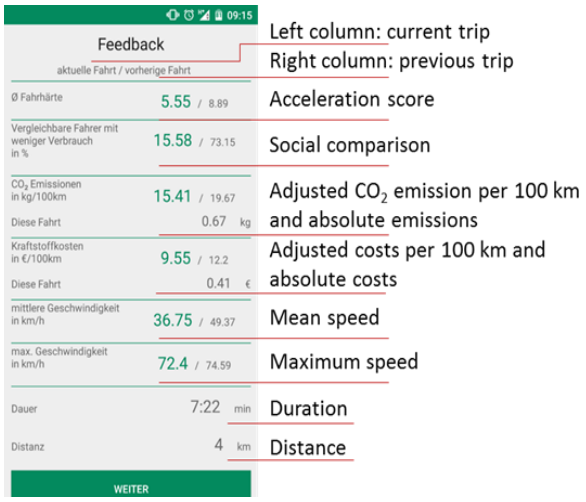


Figure 6: Screenshot of feedback (mature version) after the trip (German original with translations)

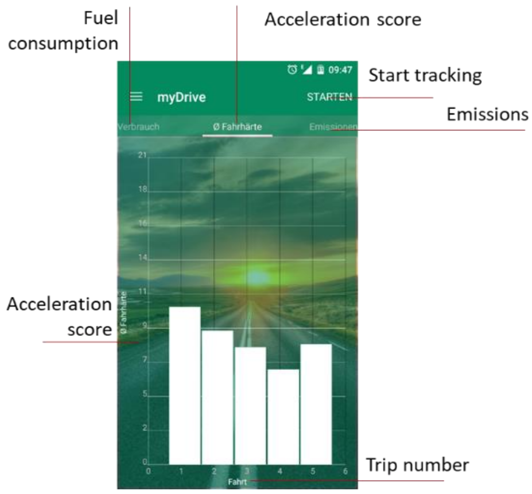


Figure 7: Screenshot of the historical comparison displaying the acceleration score within the app (German original with translations)

Field Study

We conduct a field study to gain data from real users and a real system (Venable et al. 2016). Our prototype has been available in the Google Play Store, free of charge and accessible for all interested participants. We raised awareness for the research project and our app via news reports in local newspapers, a national online news platform, a local television channel, and radio reports.

Over eight weeks, drivers used the application and uploaded their trip data. We exclude trips that contain incomplete data, such as missing speed or consumption data. The resulting data set originates from 40 participants making 604 trips (between 8 and 41 trips per participant; mean 15.1; median 13). With initial usage, all participants answer a questionnaire to reveal some information about their demographic factors, cars and car usage, and individual mindset of their environmental, cost, and social awareness. The age of the participants ranges between 21 and 60 years, and 42.7 years on average. Most are male (39), whereas merely one woman participated. Participants drive primarily in a private context (32) rather than in a business context (8).

Moreover, the majority of the participants' cars are powered by diesel engines (24) rather than gasoline engines (16). Most cars have engines with 101 to 150 hp (23), whereas three have less and 14 have more power. According to the participants' self-report using a seven-point Likert scale, they mainly have an environmentally friendly mindset (80% at least somewhat agree, level three on the seven-point scale) and care for the costs that result from driving (95% at least somewhat agree). Conversely, the relevance of social comparison is less important to them (45% at least somewhat agree).

We analyze the effectiveness of the app in lowering fuel consumption. We have information on fuel consumption (l/100km) by a trip for at least eight trips per participant. To detect potential changes over time, we compare their fuel consumption of trips one to two to their fuel consumption of trips three to eight. We assume the first two trips as a baseline as the effect of the feedback may need to evolve, and drivers need to get familiar with the app. In case the feedback does have an immediate impact when first using the app, this approach underestimates the fuel-saving effect. For now, we disregard the trips nine and following. In doing so, we mitigate concerns of a changing effect over time, depending on the total number of trips per participant. Besides, we have a sample of the same size for every participant.

We calculate the individual changes in fuel consumption for each participant. We investigate mean changes in fuel consumption for all participants. Participants decrease their consumption by about 0.35 l/100km on average (median 0.14 l/100km), which corresponds to a saving of 4% per participant on average.

Next, we investigate the effect size and the significance (significance level of 5%) of the fuel-saving effect. We calculate the difference of the mean consumption for the first two trips and the mean consumption of the six remaining trips. These differences are then tested with the statistical H_0: Fuel-saving less or equal to zero. The one-sample t-test allows to reject H_0 (p-value = 0.022). Thus, we can conclude that the reduction in fuel consumption is significant for drivers who applied the app. Although the fuel savings are not normally distributed, the t-test is justifiable as n is bigger than 30, and the t-test tends to be robust. Nevertheless, we additionally perform a Wilcoxon Signed-Rank test, leading to comparable results. Again, we see significant fuel savings (p-value = 0.025).

Subsequently, after testing for the existence of a fuel-saving effect, we also evaluate the size of the effect. Thus, we calculate the effect size, also referred to as Cohen's d. We observe an effect size of 0.21, which by convention can be considered as small effect (Table 1).

Mean of fuel saving	*Median of fuel saving*	*p-value of t-test*	*p-value of Wilcoxon Signed-Rank test*	*Cohen's d*
0.35 l/100km	0.14 l/100km	0.022	0.025	0.21

Table 1: Significance and size of the fuel-saving effect

Note: H_0 = Fuel saving is less or equal to zero

Following Karlin et al. (2015), there might be an upper limit to the time users spend dealing with feedback. Therefore, instead of comparing merely eight trips per person, we perform the analysis mentioned above also on all available trips per driver to investigate the fuel-saving effect over the full time each participant used the app. Results remain overall the same. Average fuel saving increases to 0.36 l/100km, Cohen's d slightly increases to 0.22, p-values for t-test, and Wilcoxon test are 0.022 and 0.051. This indicates a significant fuel saving effect over more than eight trips and at least for the period of active usage of the smartphone application. The increasing effect when extending the period of analysis suggests that we do not only observe an effect of initial adoption.

As a result, we found indications that our app, as an instantiation of our design principles and abstract blueprint of a mobile EDFIS, constitutes an effective solution artifact. Our proposed design artifact leads to more sustainable driving behavior. Our data indicate that behavior-related feedback, as well as feedback on rational and normative aspects of driving, encourage the majority of participants to drive more fuel-efficient. The described effect is observable for the first eight trips of each participant as well as all recorded trips up to eight weeks.

Summarizing, we show that our solution artifact meets the design requirements and DO 3-1, as we find an indication that participants decrease their fuel consumption while applying a prototypical implementation of our design artifact throughout a field study. Data analysis shows that drivers reduce their fuel consumption on average by 4% using an instantiation of the proposed design artifact. This saving is somewhat below the maximum savings of around 7% reported by van der Voort et al. (2001) and Boriboonsomsin et al. (2010) but higher than the results achieved by competing artifacts (Tulusan et al. 2012b; Dahlinger et al. 2018). Further, the instantiation creates a stronger awareness of eco-driving. It lowers fuel consumption, as it enriches the information provided by the built-in fuel consumption meter or any other feedback system that the participants of the field study had in their cars. Thus, our evaluation establishes the utility, quality, and efficacy of the design artifacts.

Discussion

Following the design science research methodology suggested by Peffers et al. (2007), this chapter presents design knowledge for a class of mobile systems that address harmful CO_2 emissions by individual traffic. Seven design principles and the resulting architecture codify knowledge on how to design a mobile EDFIS that allows drivers to adjust their fuel consumption per driven kilometer. Our architecture describes the interaction of drivers, their cars, and the feedback system. The architecture summarizes the principles of form and function of a mobile EDFIS. These principles allow measurement of driving behavior and fuel consumption to provide feedback to the driver effectively. We propose to highlight the gap between desired and actual behavior (DP 3-1) via a mobile device (DP 3-7) and to provide multiple types of feedback (DP 3-2), in particular, provide feedback on economic and environmental impact (DP 3-3) as well as normative feedback (DP 3-4). The user of an EDFIS should have the possibility to assess performance over time (DP 3-5) and receive expert advice on how to change behavior (DP 3-6).

Our design knowledge develops throughout an iterative search process. First, we analyzed the literature on eco-driving behavior, eco-feedback, and existing EDFIS artifacts that have been applied throughout various studies. Subsequently, we derive design knowledge and propose an architectural blueprint for mobile EDFIS. To demonstrate and evaluate our findings, we implement a smartphone app as a prototypical instantiation. Forty participants successfully decreased their fuel consumption by 4% on average throughout a field study making at least eight trips. In doing so, we verify the utility, effectiveness, and efficiency of the design artifacts, as well as the fulfillment of the DO 3-1 and its corresponding design requirements.

From a theoretical perspective, we present a nascent design theory with design knowledge as operational principles and architecture (level 2 contribution, according to Gregor and Hevner 2013) and a situated implementation artifact (level 1 contribution, according to Gregor and

Hevner 2013). We do not present a mid-range design theory (level 3), which constitutes a comprehensive and well-developed theory (Gregor and Hevner 2013). The derived design artifacts constitute *improvements* in the terminology of Gregor and Hevner (2013), as we propose a more effective and efficient solution for a rather well-known application domain. As to the best of our knowledge, no design theory exists, our findings present a novel and consistent nascent design theory in the field of mobile EDFIS. In conclusion, we summarize the design theory components, as suggested by Gregor and Jones (2007), in Table 2.

Design theory component	*Description*
1) Purpose and scope	Design a mobile EDFIS that supports car drivers individually in adopting a more environmentally sustainable driving behavior. First, mobile EDFIS should allow fuel savings over the same distances traveled. Second, mobile EDFIS should constitute a more effective solution than existing artifacts.
2) Constructs	The core constructs of the design artifacts are the technical system (the feedback system) and the social system (the drivers), which allow improvements while interacting. The technical system consists of a sensor layer (sensing of behavior and outcome), a feedback layer (behavior-related, rational, and normative), and a presentation layer (real-time and comparison).
3) Principles of form and function	We suggest seven design principles. These fulfill the DO 3-1 and its corresponding design requirements. Furthermore, we propose an architecture blueprint for mobile EDFIS that implements the proposed principles.
4) Artifact mutability	Sensors that become more precise and affordable, connected cars, and emerging display devices in the future enable mobile EDFIS to provide richer eco-feedback in a more distinctive way incorporating the presented design principles.

5) Testable propositions	An instantiation of a mobile EDFIS – following the architecture and the presented design principles – enables drivers to achieve higher fuel savings than with existing solutions.
6) Justificatory knowledge	The design artifacts base on existing theoretical work on eco-driving behavior, eco-feedback, and existing feedback systems. We derive our design principles from existing knowledge about mobile feedback systems, rational-choice, and norm activation, as well as eco-driving feedback studies.
7) Principles of implementation	Primarily an implementation should adhere to the suggested design principles and the architecture blueprint. Beyond these suggestions, experience from our prototypical instantiation suggests that understanding motivational factors of the users is a crucial success factor.
8) Expository instantiation	We implement a prototypical instantiation in the form of a smartphone app to demonstrate the technical feasibility and to apply the app within a field study.

Table 2: Design theory components according to Gregor and Jones (2007) for mobile EDFIS

Based on the presented design knowledge and the results from our field study (with two iterations), we find evidence that normative feedback, while controversially discussed in other contexts (Froehlich et al. 2010; Schultz et al. 2007; Allcott 2011), plays a vital role with regards to eco-driving. The first version of the prototype, without normative feedback, did not lead to adjusted driving behavior, while the inclusion of normative feedback in the second iteration led to a significant change in driving behavior. Although only about half of the probands per iteration reported that social comparison is at least somewhat important to them[6], the extension of the prototype with normative feedback (DP 3-4) leads to significant improvements in reported fuel consumption. One possible explanation is that social comparison motivates drivers to achieve higher fuel savings as compared to other drivers (Magana and Organero 2013; Ecker et al. 2011). Another explanation is that feedback about the behavior of others highlights the

[6] The participants' characteristics of both groups are comparable as a Chi-squared test of homogeneity is not significant for all reported characteristics, in particular regarding the importance of social comparison.

need for more pro-environmental behavior and motivates eco-driving behavior as an accepted and desirable behavior (cf. focus theory of normative conduct, see Cialdini et al. (1990)). Overall, giving normative feedback in the context of eco-driving is a key finding of this study. Interestingly, we could not observe any significant correlation between the self-reported importance of social comparison and achieved fuel savings. This finding indicates that the participants of the field study were accessible for normative feedback independently, whether the social comparison is essential to them or not.

Systems building on our derived design knowledge and fulfilling the DO 3-1 and corresponding design requirements can have a share in fighting climate change. From a practical perspective, the design knowledge is helpful for car drivers, who can contribute to ecological sustainability, reduce their fuel expenses, and contribute to more environmentally friendly societies. Furthermore, these findings are beneficial for companies in various cases. Firstly, automotive manufacturers and their suppliers can offer more sustainable and environmentally friendly products. EDFIS enable a more fuel-efficient use of vehicles while manufacturers are working on alternative powertrains. Secondly, third parties are entering the market and offering solutions for connected vehicles in the IoT. IoT-based smart vehicle services can also be used as EDFIS and thus contribute to a more sustainable way of life (see also Chapter 3.2). In times of Friday-for-Future demonstrations, it is possible that an attractive market for services that enable a more sustainable lifestyle will emerge. Thirdly, companies have been striving that corporate car drivers have an efficient driving style and thus save fuel costs. Various studies have been conducted with field staff (Tulusan et al. 2012b) or mail-van drivers (Siero et al. 1989), for example, to investigate the impact of feedback on fuel consumption. EDFIS can, therefore, generate savings without offering financial incentives to employees.

Naturally, our findings are limited, as the results cannot be considered complete or universal. We find an indication that the derived design principles help to decrease fuel consumption on average and for the majority of drivers. However, these findings are limited by a few factors. First, the results do not explain which feedback works best for a specific individual. It is possible that by combining different types of feedback, the driver receives the appropriate feedback and ignores the other feedback types. Personality has an impact on driving behavior (e.g., Shahab et al. 2013) and also on the effect of feedback (Ilgen et al. 1979; Brouwer et al. 2015). Thus, future research should further investigate the dependency of different feedback types and personality traits and derive implications refining our design artifacts. Second, we solely examine the effect of anonymous social comparison, whereas comparison with known persons can have other impacts (e.g., Schultz et al. 2007; Foster et al. 2010). A context-dependent detailing of a specific implementation that addresses driving behavior, to decrease CO_2 emissions, may provide additional principles. Third, we have investigated the effect of feedback based on the first two trips as a baseline compared to the following trips, and the maximum usage was only eight weeks. Further, in our field study, we do not have a control group, not obtaining feedback. It might be the case that we recruited individuals specifically eager to change their driving behavior (self-selection bias), who then did so more or less independently of the EDFIS specific functionality. However, having a control group report their fuel consumption without triggering reflection on driving behavior and, thus, potentially changing driving behavior was not possible, and a certain willingness to adopt and use an EDFIS will always be a prerequisite for the system‘s success. Further research is needed to verify the effectiveness of our derived design knowledge in different settings and with different instantiations. Besides, we have not focused on safety-critical factors. Feedback mustn't distract the driver. Safety-critical factors may influence the design of a system (Jamson et al. 2015) and should, therefore, be investigated in future research. Finally, we do not consider different trip profiles. That is, we do not differentiate between city streets, highways, speedways, or various covered distances. However, different trip profiles imply distinct fuel consumption and savings, as several authors illustrate (e.g., Boriboonsomsin et al. 2010).

Future research should further analyze our design artifacts, as well as extend and refine them. Researchers should investigate diverging samples of drivers and various human characteristics (e.g., values, preferences, demographic factors) to examine whether our design recommendation is efficient for different types of drivers. Future work might explore the transferability of knowledge for mobile EDFIS to EDFIS in general. Future work also should investigate similarities with, and differences to, feedback systems in other domains beyond eco-driving or smart driving systems that may become more important in the future, such as eco-feedback systems in smart homes.

In conclusion, we address the design of mobile EDFIS to improve driving behavior to decrease fuel consumption. Harmful emissions at the hand of individual traffic constitute a considerable percentage of global GHG emissions and are associated with severe social consequences. Thus, better driving behavior contributes to the global challenge of achieving a more sustainable lifestyle.

3.2. Analyzing Environmentally Friendly Behavior using IoT-Data – Insights into Real-World Driving Behavior

The previous Chapter 3.1 demonstrates that providing sufficient eco-feedback results in fuel savings. That is, eco-feedback must affect driving behavior, which has a direct impact on fuel consumption (Lárusdóttir and Ulfarsson 2014; Ericsson 2001). Some studies already considered how distinct driving variables change throughout the application of eco-feedback. For instance, Martin et al. (2013) analyze the effect of eco-feedback on fuel consumption, as well as acceleration, deceleration, and average speed. However, research indicates that eco-driving strategies could lead to much higher fuel savings between 5-30% (Sivak and Schoettle 2012) as compared to the fuel savings in eco-feedback studies. Hence, it is important to understand how driving behavior changes while providing eco-feedback to further develop advanced user assistance systems (Maedche et al. 2016) that encourage and enable drivers to adopt eco-friendly driving strategies (Saboohi and Farzaneh 2009). Yet, to date, we lack knowledge of how driving behavior changes due to eco-feedback. A prime reason is the traditional lack of data. Average fuel consumption over longer distances and time spans can be measured rather easily. However, until recently, it was practically impossible to obtain detailed data on driving maneuvers on scale.

Nowadays, cars' built-in sensors and the standardized OBD-II interface allow access to rich data. IoT-based smart vehicle services collect comprehensive and detailed data on vehicle activities and driving behavior, such as acceleration, speed, and revolutions per minute (RPM). From this data, we can expect a better understanding of the effect of eco-feedback on driving behavior. Crucially, the data reflects authentic and unbiased real-world driving behavior rather than data biased by retrospective self-assessments or behavioral change due to the awareness of taking part in a research project.

The purpose of our research is to gain a better understanding of the effect of eco-feedback on driving behavior by means of IoT-data. For this reason, our research method consists of five steps. First, real-world driving-data was collected over ten weeks from 5,676 users of an IoT-based smart vehicle service. The users were selected randomly from the service's customers. We split the data into two halves and conduct an exploratory factor analysis (EFA) on the first half to derive factors characterizing driving behavior (step 2). Subsequently, we validated factors for driving behavior by means of a confirmatory factor analysis (CFA) on the second half (step 3). Thus, each analysis uses data from 2,838 users. Next, we developed theoretically deduced hypotheses on the effect of eco-feedback on our factors of driving behavior (step 4). To test these, we collected real-world driving data from another 495 users who are both customers of the IoT-based service and customers of an insurance company and, therefore, qualified for the eco-feedback feature newly introduced by the two companies together. After the first four weeks, the eco-feedback was launched in order to provide eco-feedback to the user group via the use of a mobile app (an integral part of the IoT-based service). Subsequently, we recorded the driving data for another six weeks. Using hypothesis tests, we derive first insights about the effect of eco-feedback on driving behavior (step 5).

Therefore, the remainder of this chapter is structured as follows: First, we discuss the theoretical background. Next, we elaborate on our methodology, data set, and the study design followed by the results of the factor analyses. Subsequently, we derive initial hypotheses and gain first insights. Finally, we discuss our findings.

Theoretical Background

As already shown throughout Chapter 3.1, driving behavior affects vehicle fuel consumption. Furthermore, various studies found evidence that eco-feedback is a suitable approach to achieve fuel savings. Hence, feedback must have an effect on one's driving behavior that leads to a decrease in fuel consumption (Ericsson 2001; Lárusdóttir and Ulfarsson 2014). However, little has been done to understand the effect of eco-feedback on driving behavior. Given the reduction of fuel consumption upon eco-feedback, it can be assumed that drivers change some aspects of their driving behavior based on eco-feedback. However, it seems that optimal driving strategies can save more fuel than the results of the eco-feedback studies show (cf. Sivak and Schoettle 2012; Tulusan et al. 2012b; Boriboonsomsin et al. 2010).

An explanation is that the effect of eco-feedback does not sufficiently address all factors of driving behavior. For instance, Martin et al. (2013) show that eco-feedback on fuel consumption and CO_2 emissions reduces fast driving and favors slow driving. Vagg et al. (2013) provide rather behavior-specific feedback on the uneconomical power demands of the engine. They aim to reduce acceleration and early upshifting. Results show successful improvements in relation to these two driving factors. Graham et al. (2011) provide feedback on car usage and achieve significant reductions in annual mileage.

Summarizing, eco-driving has an impact on fuel consumption and, thus, GHG emissions. Numerous scientific studies have shown that feedback is a possible approach to motivating ecological driving behavior and thereby saving fuel. However, there is a lack of research investigating which factors from data of an IoT-based smart vehicle service describe driving behavior and how these are affected by eco-feedback.

Method

Design of the Eco-Feedback

In addition to our study design (as stated in the introduction), the smart vehicle service provides eco-feedback to 495 drivers, delivered via a mobile app. At the end of each trip, the service sends an eco-score between 0 and 100% to the participant's mobile device. The total score is composed of four separate scores, each ranging from 0 to 25. Together the scores are adding up to a maximum of 100. The four scores reflect an average acceleration in comparison to other users of the IoT-service; average deceleration in comparison to other users of the IoT-service; trip length (short journeys (<5 km) by car are classified as less environmentally friendly); and the proportion of the trip driven at speeds in excess of 130 km/h (which is a reasonable threshold in Germany, as there is no general speed limit). Participants receive a push notification at the end of each trip (Figure 8, bottom) with the score and some details about length, duration, speed, and acceleration of the completed trip. In addition, participants are able to view their latest score, as well as the average score, when they open the mobile app (Figure 8, top).

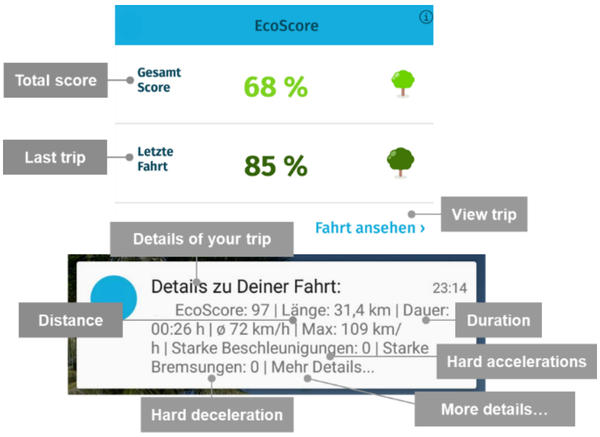


Figure 8: Visualization of the eco-score in the app (top) and as a push notification (bottom)

Variables in the Data Sets

The service records driving data via the OBD-II interface. Recording data enables the service to offer its users insights into their cars (e.g., battery level, location of the car), and location-based services (e.g., discounts at selected petrol stations), via their mobile devices. Since our data was recorded before the inception of this study, users were not aware of their inclusion, and thus our results are unbiased. To preserve privacy, and in accordance with the service's privacy policy, we do not know the identities of users, nor do we have access to any location data. All included users have agreed to the anonymous use of their data in advance by accepting the privacy policy of the IoT-based smart vehicle service. From the smart vehicle service, we draw the following available data for each short stretch of the way of each trip: speed (in km/h), acceleration, deceleration (both in g), RPM, driving time, standing time (both in seconds), driven distance (in meter), throttle (in degree), and engine coolant temperature (in °C).

Unfortunately, the service does not record which gears are engaged. Instead, we use the RPM as an approximation of the gear-shifting behavior. The variables for acceleration and deceleration exhibit outliers. Some observations exceed or fall below plausible values and, thus, are capped. We computed the 99%-quantile for acceleration and deceleration and capped observations to eliminate outliers. As the sensing device delivers data continuously, there was no need to handle missing data. Following this data cleansing, we aggregated the single values of each variable for a whole week in order to enable a meaningful comparison of a person's driving behavior in regular time intervals. Thus, the aggregation reduces inaccuracies, which may result, for instance, from differences in driving behavior on working and leisure days, as well as effects of chance. The aggregation is done by computing several of the following summary statistics: minimum, maximum, sum, average, variance, and 95%-quantile.

Results

Identification of Factors Characterizing Driving Behavior

To assess the effect of feedback on driving behavior, we need to clarify the measurable factors of driving behavior in our IoT data. Consequently, we conduct an EFA on the driving data of 2,838 drivers in R (Revelle and Revelle n.d.). In total, the data comprises of 36 variables describing the driving behavior (different weekly aggregations of the aforementioned variables). Thirty-five variables fulfill the Kaiser-Meyer-Olkin criterion with a measure of sampling adequacy (MSA) above 0.5 and thus are further analyzed. Overall the MSA is 0.79, which indicates the data is suitable for conducting an EFA. The same applies to Bartlett's test of homogeneity, which is significant (p-value < 0.001). A parallel analysis (Horn 1965) initially suggests ten factors. After oblimin rotation, we excluded items with a major loading below 0.4. As a result, one variable (average engine coolant temperature) is dropped. Furthermore, no item substantially loads on the tenth factor, which is then removed from the analysis. Rerunning the factor analysis with nine factors results in the loadings displayed in Table 3. Factor nine is a subset of factor four, and due to the higher loadings and eigenvalue of factor four ignored. The result of the EFA is eight factors that describe aspects of strategic and operational driving behavior.

To ensure our results will be useful, we need to be able to interpret our factors. If a single factor cannot be interpreted, the usefulness of the results is limited. In the following, we describe the eight factors identified, which serve as our understanding of driving behavior. The first factor can be interpreted as *acceleration behavior*. It includes the average, maximum, 95%-quantile, and variance of acceleration and deceleration. The second factor is the *total driving per week*, composed of the total distance, time in the car, and the number of trips per week. The third factor can be interpreted as the *speed behavior*, encompassing the average and maximum speed as well as the variance of speed. The fourth factor is related to the first factor, *acceleration behavior*, with the difference that it covers the number and frequency of extreme events of acceleration or deceleration. The factor can be interpreted as *extreme acceleration behavior*. The smart vehicle service provider defines the thresholds for those events. The fifth factor reflects the *average driving per week* in the sense of average trip distance and average trip duration. The sixth factor can be interpreted as *minimum driving per week,* which reflects the shortest trip distance and duration per week. The seventh factor is *RPM behavior* consisting of the average, maximum, and variation of the RPM of the engine. Last is the *throttle behavior*, reflected by average, maximum, and variation of the throttle position.

Most factors include at least one loading greater than 0.7, but one, the throttle factor, has two indicators slightly missing that threshold. These eight factors cover 66% of the variance in the data (cf. Table 4). Cronbach's alpha for our factors is mostly excellent. Only the throttle factor does not reach 0.7 and, thus, is removed for further analysis.

	Factor								
Original Parameter	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Average acceleration	0.73								
Maximum acceleration	0.84								
95%-quantile acceleration	0.76								
Variance acceleration	0.75								
Average deceleration	0.79								
Maximum deceleration	0.89								
95%-quantile deceleration	0.89								
Variance deceleration	0.89								
# trips		0.85			-0.43				
Sum distance		0.64							
Sum duration		0.91							
Sum driving duration		0.83							
Sum standing duration		0.93							
# long stops (>3min)		0.69							
# stops		0.51							
Average speed of trip averages			0.94						
Average speed			0.64						
Average speed when driving			0.93						
Maximum speed			0.70						
Variance speed			0.70						
# hard accelerations				0.75					
Hard accelerations per km				0.85					

	1	2	3	4	5	6	7	8	9
# hard decelerations				0.47					0.45
Hard decelerations per km				0.53					0.47
Average trip distance					0.70				
Average trip duration					0.86				
Minimum trip distance						0.90			
Minimum trip duration						0.92			
Average RPM							0.66		
Maximum RPM							0.85		
Variance RPM							0.68		
Average throttle								0.68	
Maximum throttle								0.67	
Variance throttle								0.55	

Table 3: Rotated factor matrix from EFA (with loadings ≥ 0.4)

Note: Factors can be named as follows: (1) acceleration behavior, (2) total driving per week, (3) speed behavior, (4) extreme acceleration behavior, (5) average driving per week, (6) minimum driving per week, (7) RPM behavior, (8) throttle behavior, (9) disregarded as a subset of factor (4)

For some factors like RPM, it might appear straightforward that average, maximum, and variance are highly correlated, although technically, this is not necessarily the case but a result. For other factors, the structure is less straightforward; examples are the combination of acceleration and deceleration in the first factor but the separation of factors one and four.

Overall, our factor model seems to be consistent with the factors identified by Ericsson (2001). However, due to different availability of data, our factor model encompasses three factors of strategic driving behavior (total, average and minimum driving per week) and four factors related to operational driving behavior (acceleration, extreme acceleration, speed, and RPM behavior), while Ericsson (2001)'s factor model solely consists of factors related to operational driving behavior. Therefore our factor model is more comprehensive.

	Factor								
Statistics	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Proportion Variance	0.16	0.13	0.11	0.06	0.05	0.05	0.05	0.04	0.02
Cumulative Variance	0.16	0.29	0.41	0.47	0.52	0.57	0.62	0.66	0.68
Cronbach's Alpha	0.96	0.91	0.91	0.83	0.97	0.91	0.75	0.51	-

Table 4: Descriptive statistics of factors in EFA

Note: Factor numbering as in Table 3, Cronbach's Alpha should be >0.7

Validation of factors characterizing driving behavior

Following the development of our model for driving behavior, we validate our factors for driving behavior by applying a CFA on an independent data set from the smart vehicle service provider. The data set results from a new set of 2,838 users. The CFA shows good values (>0.7) in terms of Cronbach's Alpha (Cα) for all seven factors. Values of 0.6 regarding the composite reliability (CR) and 0.5 for average variance extracted (AVE) can be seen as good measurement quality (Bagozzi and Yi 1988). All seven factors did meet these thresholds.

We check the discriminant validity for the seven factors by using the Fornell-Larcker criterion that a factor's AVE should be higher than its squared correlation with every other factor (Fornell and Larcker 1981). Table 5 summarizes the results. We can assume discriminant validity for all seven factors of driving behavior.

In summary, we can state that we have found – based on the data provided by a smart vehicle service provider – a valid factor structure for assessing driving behavior.

				Fornell-Larcker criterion						
Factor	Cα	CR	(1)	(4)	(3)	(7)	(6)	(5)	(2)
(1) Acceleration behavior	0.96	0.96	**0.77**						
(4) Extreme acceleration behavior	0.79	0.90	0.47	**0.82**					
(3) Speed behavior	0.91	0.84	-0.14	-0.04	**0.53**				
(7) RPM behavior	0.91	0.92	0.11	0.12	0.24	**0.79**			
(6) Minimum driving per week	0.95	0.98	-0.06	-0.03	0.29	-0.05	**0.95**		
(5) Average driving per week	0.97	0.98	-0.29	-0.04	0.67	0.11	0.42	**0.97**	
(2) Total driving per week	0.90	0.93	-0.12	0.11	0.18	0.13	-0.15	0.37	**0.74**

Table 5: Descriptive statistics of CFA

Note: Factor numbering as in Table 3; For Fornell-Larcker criterion, diagonal elements represent AVE and off-diagonal elements correlations.
Good measurement quality if Cronbach's Alpha >0.7, Composite-reliability >0.6, AVE >0.5, and Fornell-Larcker criterion met, i.e., AVE is higher than the squared correlation with any other factor

Developing hypotheses for driving behavior affected by feedback

Having identified seven factors that describe driving behavior on the basis of our IoT-data set, in the following, we introduce hypotheses on how the eco-feedback could affect driving behavior.

Martin et al. (2013) find evidence that both acceleration and deceleration are likely to improve if drivers receive eco-feedback on their driving behavior. Especially forceful acceleration increases fuel consumption and is therefore likely to be decreased by knowledgeable and eco-conscious drivers (Ericsson 2001; Lárusdóttir and Ulfarsson 2014). Deceleration, in contrast, has no direct impact on fuel consumption but causes acceleration back up to speed (Lárusdóttir and Ulfarsson 2014).

Furthermore, the eco-feedback of the IoT-based smart vehicle service punishes hard acceleration and deceleration by a deterioration of the eco-score. The driver, in addition, receives the number of hard acceleration and deceleration events in the trip summary. Thus, we assume that acceleration and deceleration will improve when providing feedback to the driver, which is reflected in two factors.

H 3.2.1 Providing eco-feedback decreases the factor "acceleration behavior"

H 3.2.2 Providing eco-feedback decreases the factor "extreme acceleration behavior"

In addition to changes in acceleration and deceleration behavior, Martin et al. (2013) also find evidence that eco-feedback reduces speed behavior. Speed impacts fuel consumption resulting in fuel savings, especially when driving at a moderate speed (between 50 and 70 km/h) (Lárusdóttir and Ulfarsson 2014; Ericsson 2001). Accordingly, we assume that our feedback has a reducing effect on speed.

H 3.2.3 Providing eco-feedback decreases the factor "speed behavior"

Another factor of driving behavior from our factor analysis is the RPM, which relates to speed and gears engaged. Eco-friendly driving requires early upshifts and, therefore, low RPM (Saboohi and Farzaneh 2009). However, the eco-feedback of the IoT-based smart vehicle service does not address either the RPM or the shifting behavior. For this reason, we assume that this factor will not change, although an eco-friendly strategy would require low RPM.

H 3.2.4 Providing eco-feedback does not affect the factor "RPM behavior"

Eco-feedback can address driving behavior on a rather strategic level as well. Graham et al. (2011) use eco-feedback successfully to reduce the number of rides and, therefore, the overall mileage. The study reflects the participants what environmental (CO_2) and financial savings have been achieved by not using the car. Eco-feedback from the IoT-based smart vehicle service reduces the eco-score when the vehicle is used for short distances, reflecting non-ecological use. Accordingly, we expect that short distances are avoided and, therefore, the remaining trips become longer on average.

H 3.2.5 Providing eco-feedback increases the factor "minimum driving per week"

H 3.2.6 Providing eco-feedback increases the factor "average driving per week"

The last factor of driving behavior in our study is the factor *total driving per week*. We assume that the distance covered decreases because short distances are avoided (see H 3.2.5 and H

3.2.6). In addition, the eco-feedback could create awareness for every trip taken (Graham et al. 2011). Thus, we hypothesize that users may leave their cars and use alternative means of transport, leading to less overall driving per week.

H 3.2.7 Providing eco-feedback decreases the factor "total driving per week"

Analysis of the effect of eco-feedback

Finally, to assess the effect eco-feedback has on driving behavior, we conduct our analysis on a third data set consisting of 495 drivers. For every driver, we have baseline driving data of four weeks. The eco-feedback was launched during week 5. To ensure the baseline and the treatment phase are not mixed, this week is removed from the data set. Following the launch of the eco-feedback, the data set comprises of six weeks of data per driver. As not every car was used every week, the baseline record per driver is 3.79 weeks and 5.37 weeks for the treatment phase on average. Based on the factors we identified, we calculate the factors of driving behavior per week per driver. We assess the effect of whether and in which direction feedback influences driving behavior based on the aggregated driving factors. Consequently, the factor scores per driver are averaged before and after the start of the feedback. First, we check whether the aggregated factor scores are normally distributed to choose the appropriate test. Consequently, we perform the Shapiro-Wilk test of normality.

The results suggest that factor scores are not normally distributed across the drivers. As a consequence, we compare the factor scores before and after the launch of the feedback applying the Wilcoxon signed-rank test. For H 3.2.4, where we assume no effect, we perform a two-sided test. For all other hypotheses, we perform one-sided tests. In addition, we check not only the presence and direction of an effect but also the effect size (Cohen's d). The effect size is calculated by subtracting the mean of the experimental group (here: after application of eco-feedback) from the mean of the control group (here: before the application of the eco-feedback) and dividing the difference by the standard deviation of the data. Thus, a negative value of Cohen's d indicates a decrease in the respective factor due to the display of feedback. A positive value vice versa indicates an increase in the factor score. An absolute value of the effect size of 0.2 is termed a small effect (Coe 2002). The results of the Wilcoxon signed-rank test and the effect size are depicted in Table 6.

Our data supports H 3.2.1 and H 3.2.2. We conclude that in terms of acceleration behavior as well as extreme acceleration behavior, feedback does have a desired, significant, and meaningful (small) effect. Drivers seem to accelerate and decelerate more carefully.

In terms of speed and RPM, we find no significant effects of the eco-feedback. Consequently, we reject H 3.2.3, whereas H 3.2.4 is supported. However, while the average speed slightly increases from 31.51 km/h to 31.76 km/h, the average maximum speed slightly decreases from 135.51 km/h to 134.69 km/h, which appears to be favorable as especially high speeds cause higher fuel consumption.

	Factor	p-value of Wilcoxon signed-rank test		effect size	Hypothesis	Result
(1)	Acceleration behavior	<0.001	***	-0.19	H 3.2.1: Decrease through feedback	Support
(4)	Extreme acceleration behavior	<0.001	***	-0.18	H 3.2.2: Decrease through feedback	Support
(3)	Speed behavior	0.949		0.06	H 3.2.3: Decrease through feedback	Reject
(7)	RPM behavior	0.203		-0.03	H 3.2.4: No effect through feedback	Support
(6)	Minimum driving per week	0.016	*	0.07	H 3.2.5: Increase through feedback	Support
(5)	Average driving per week	0.015	*	0.12	H 3.2.6: Increase through feedback	Support
(2)	Total driving per week	0.025	*	-0.07	H 3.2.7: Decrease through feedback	Support

Table 6: Summary of effect of eco-feedback on driving behavior

Note: *** p-value <0.1%, ** p-value <1%, * p-value <5%

The factors for minimum and average driving per week are significantly influenced by the application of the eco-feedback in that the minimum and average trip length and duration increase. This supports H 3.2.5 and H 3.2.6. It seems as if our assumption was correct that short distances are avoided, and thus the average trip becomes longer. Both the minimal trip distance and the average distance per trip increase from 459 m to 606 m and 11.24 km to 12.02 km, respectively. The effect is statistically significant, considering the effect size but not substantial for both factors.

Total driving per week is significantly influenced by feedback, which supports H 3.2.7. However, the magnitude of the effect seems to be negligible. Nevertheless, we observe that, for example, the number of trips per week decreases by 0.54 trips per user of the IoT-based smart vehicle service.

Discussion

In this chapter, we derive seven factors from data of an IoT-based smart vehicle service, which describe strategic and operational driving behavior. Furthermore, we investigate how eco-feedback affects driving behavior. We show which aspects of driving behavior are positively influenced due to the application of eco-feedback in a real-world scenario.

A strength of our study is the large real-world data set exceeding previous studies, arguably leading to more generalizable results. In addition, our data is not biased by the Hawthorne effect (Rice 1982), as participants did not know their driving behavior was investigated in the context of the eco-feedback. We, thus, can assume the observed effects are due to the intrinsic motivation of the participants and not due to the fact that they were asked to participate in a study. Even if the measured effects are only small or almost negligible, if the lever is big enough, these effects still can make an important impact. With regard to climate change, almost any effort is important, and even small steps can contribute to making transportation more sustainable.

The theoretical contribution of this chapter is a comprehensive factor model explaining driving behavior on a strategic and operational level. In comparison to Ericsson (2001), the variables underlying our factor model do not need additional specific data collection equipment as the respective data can be obtained via the OBD-II-interface and, thus, is widely applicable, especially as IoT-solutions push into the market for additional driving features. Further, we have shed light on how drivers adjust their driving behavior based on eco-feedback, which reflects how drivers expect to drive more sustainably. While we observe the strongest effects with regards to operational driving behavior (acceleration behavior and extreme acceleration behavior), the effects on strategic driving behavior (average, minimum and total driving per week) seem to be smaller and, thus, eco-feedback seems to be less effective with regards to this concern. Conclusively, eco-feedback alone may not be sufficient to grasp the full potential for more eco-friendly driving. For example, advanced user assistance systems (Maedche et al. 2016) may supplement feedback with other approaches like goal-setting, rewards (Froehlich et al. 2010), or personalized recommendations (Vagg et al. 2013) to improve previously unaddressed behavior along the path towards autonomous vehicles.

Besides the theoretical contribution, this chapter offers managerial implications. Our findings are relevant with regard to designing future feedback systems in the automotive sector. We derive first insights that allow car manufacturers, insurance companies, as well as third-party applications to tailor feedback to make it more effective. Practitioners and researchers may build on our findings to gain a better understanding of how to design advanced user assistance systems (Maedche et al. 2016) to reduce the emission of GHG. In doing so, IS can contribute to a more sustainable lifestyle and help to reduce harmful GHG emissions. Moreover, lower GHG emissions result from reduced fuel consumption, which leads to lower costs. Thus, feedback may help companies, especially logistics companies, to save money as their daily business consists of many vehicles and drivers. Further, eco-friendly driving tends to go along with safe driving, which is the reason why a car insurance company sponsored the development of the eco-feedback functionality in the IoT-service. Hence, the benefits of IoT-based eco-feedback may go beyond the positive effect on environmental sustainability.

Despite the rigor of our study, our findings are subject to some limitations. We provided the eco-feedback only within a mobile app. We could not ensure that all participants regularly checked their eco-feedback or truly received the push notifications. Furthermore, we cannot be certain that the cars involved in the study were driven only by our participants. Thus, the presumed effects of eco-feedback on driving behavior might, in fact, be stronger when the feedback would be more salient in the car. Our data set is limited by the variables that the smart vehicle service has chosen to measure and disclose to us. As a consequence, our factor model does lack certain variables that are not measured or disclosed by the smart vehicle service provider, like gear-shifting, type of road, or the actual fuel consumption. Additional variables could enhance the factor model and further investigate and improve the effectiveness of feedback. Our sample was restricted to customers of the service, which implies a limitation to Germany and possibly a self-selection bias as customers are presumably more interested in vehicles and potentially care about their driving style. We only considered the effect of eco-feedback in the short term as our data set only contains information about the variables for ten weeks in total. Finally, the analysis of the effect of eco-feedback does not include a control group and, thus, might be affected by unmeasured or uncontrolled external conditions, i.e., changes in weather between the period prior and after the launch of the eco-feedback. We are in the process of obtaining data for a control group. Nevertheless, the findings offer promising first insights and provide a starting point for future research.

Based on our factor model, more sophisticated analyses are conceivable, which could consider that, for example, speed or RPM is no linear function in terms of fuel consumption and the effectiveness of feedback. However, not only the analysis could be extended, but also the model itself. Hence, the measurement of additional variables could describe driving behavior in more detail. In addition, future work could focus on specific groups of drivers, selected on the basis of either similar driving behavior or personal factors. Personal factors could be of relevance in this field as Lewin's equation states that behavior – here driving behavior – is a function of the environment (here: among other influences, the provided feedback) and the person, respectively personal factors, which are not investigated in this study (Lewin 1936). This will allow further investigations into the effects of eco-feedback on specific sub-groups and will, thus, enable more customized and effective feedback in a real-world setting. In addition, driving patterns could be used to evaluate different types of feedback in order to increase impact, as the feedback applied in our study presumably influenced participants with environmental awareness. Finally, future research might investigate whether a person's (operational) driving behavior is unique – like a fingerprint – and, if so, whether it may, for example, be used to prevent insurance fraud.

In summary, we believe that data from IoT-based smart vehicle services offer a promising opportunity to better understand the effect of feedback and to make feedback and advanced user assistance systems even more efficient.

4. Design Knowledge from a Sociocentric Perspective

Sustainable development is a complex and difficult task that requires individual and collective actions (Messerli et al. 2019). Henkel and Kranz (2018) argue that pro-environmental behavior is crucial to achieving sustainability on an individual, organizational, and societal level. However, due to their findings, more should be done concerning belief formation (Henkel and Kranz 2018). Scientific findings and appeals are counteracted by non-environmental interest groups in that false information are deliberately disseminated (van der Linden et al. 2017). For this reason, this chapter aims at developing IS design knowledge to avoid negative effects from false information and, thus, contributes to pro-environmental belief formation. Therefore, this chapter takes a sociocentric perspective. The aim is to generate knowledge on how social structures and processes influence the design of technologies. Chapter 4.1 provides a detailed introduction to the threat of fake news, especially in social media. Chapters 4.2 and 4.3 provide novel insights to improve fake news detection as well as fake news reporting. Major parts of Chapter 4 conform with Gimpel et al. (2020a) and Gimpel et al. (2020b).

4.1. Introduction to the Threat of Fake News in Social Media

The increasing importance of social media in daily life has changed the way news is consumed and produced. More and more people tend to search and consume news from social media instead of using the services of traditional news organizations (Gottfried and Shearer 2016; Newman et al. 2017; Shu et al. 2017). Likewise, news production changed as an important part of social media is the ability to create and share content (Kaplan and Haenlein 2010). This ability leads to a lack of authorities controlling information dissemination on social media (Shao et al. 2016). Information verification — traditionally part of journalistic efforts — can no longer be ensured, making social media vulnerable to the spread of fake news (Rubin et al. 2015).

Fake news articles “are intentionally and verifiably false” (Allcott and Gentzkow 2017, p. 4). They became more prevalent in recent years (Shao et al. 2016) and spread “significantly farther, faster, deeper, and more broadly” than true news (Vosoughi et al. 2018, p. 1146). This trend also increases negative effects, such as leading individuals to believe inaccurate information and influence their opinions and actions (Allcott and Gentzkow 2017). Fake news makes it difficult to distinguish between truthful news and false ones (Allcott and Gentzkow 2017), which increases the state of distrust and confusion (Chatfield et al. 2017; Shu et al. 2017). The widespread dissemination of fake news could lead to a reduction in the demand for authentic information (Allcott and Gentzkow 2017).

Fake news is a political and social threat. Therefore, there is a need for countermeasures against it. With respect to fake news, “a nascent body of IS research is emerging” (Bernard et al. 2019, p. 1), for example, highlighting source-ratings (Kim et al. 2019) as countermeasures. The doctoral thesis at hand includes findings to add to the body of knowledge.

Fake News

Fake news in itself is not a new phenomenon. In the past, various groups of people, such as advertisers, political activists, and religious zealots, have used different forms of fake news to influence public opinion or to spread propaganda (Mustafaraj and Metaxas 2017; Shu et al. 2017). Two key characteristics make news stories fake news: inauthenticity and deception. On the one hand, fake news contains information that is verifiably false but is nevertheless deliberately disseminated (Shu et al. 2017). On the other hand, fake news has the potential to mislead consumers, whereby false information is interpreted as truthful (Allcott and Gentzkow 2017; Mustafaraj and Metaxas 2017). Fake news appears in the form of text, images, audio files, or videos that are edited or deliberately taken out of context (Mustafaraj and Metaxas 2017; Shu et al. 2017).

Besides the different forms of fake news, there are also different motivations for individuals and organized groups to produce and distribute false information. First, the use of fake news, such as satirical content or hoaxes, serves as a source of entertainment. Second, due to their eye-catching and sensation-oriented design, fake news is well suited to induce readers to visit the original source (Rubin et al. 2015; Shu et al. 2017) and thereby generate a financial profit through advertising (Allcott and Gentzkow 2017). Third, fake news can be used to influence opinion. The aim may be to create an advantage for the favored political candidate (Allcott and Gentzkow 2017; Mustafaraj and Metaxas 2017; Shu et al. 2017; Warzel and Mac 2018; Balmas 2012) or to influence public opinion on a political or social issue. For example, the fake news spread that refugees raped a 13-year old girl in Germany. Based on this false information, hundreds of people started to protest against the attack resulting in a diplomatic dispute between Germany and Russia (Connolly et al. 2016). Therefore, the main danger of fake news lies in the fact that people do not recognize the false information and the associated manipulation. This could lead to an unjustified change of voting behavior or political or social attitudes resulting in a serious negative impact on individuals and society.

Even correcting misinformation does not necessarily change people's beliefs (Nyhan and Reifler 2010; Flynn et al. 2017). Furthermore, even when an individual believes the correction, the misinformation may persist. Any repetition of misinformation can have negative consequences (Greenhill and Oppenheim 2017) due to familiarity and fluency biases. Even in the context of refuting it. The more people hear a story, the more familiar it becomes, and the more likely they are to believe it as true (Hasher et al. 1977; Schwarz et al. 2007; Pennycook et al. 2018). As a result, exposure to misinformation can have long-term effects, while corrections may be short-lived. Hence, limiting the spread and effect of fake news is important.

Fake News in Social Media

Although fake news has already been used in the past, its significance and dissemination have recently gained in importance. Social media, in particular, provides a platform for the effective spread of fake news. First, social media offers a large audience (Facebook 2019; Twitter International Company 2018). For substantial parts of this audience, social media is an important source of news (Gottfried and Shearer 2016; Newman et al. 2017), which makes its users highly vulnerable to fake news. Second, the barriers to enter social media are low (Allcott and Gentzkow 2017). The lack of control during account creation allows an individual to remain almost anonymous. This reduces the inhibition threshold for the quality of published information since the author hardly has to fear any consequences. The low effort involved in setting up an account also enables the creation of machine-controlled user accounts, such as social bots (Ferrara et al. 2016). As a result, about 19 million machine-controlled accounts were active on Twitter during the U.S. presidential election 2016, spreading information about the candidates (Oxford Internet Institute, University of Oxford 2016). Third, transmission speed and reach through social media are higher than ever before. Since exchanging information is an essential part of social media (Kaplan and Haenlein 2010), it is easier to discuss or further share both real and fake news stories (Shu et al. 2017).

About 25% of U.S. adults reported that they shared fake news deliberately or unknowingly (Barthel et al. 2016). Therefore, information can quickly reach global access in a viral fashion (Shao et al. 2016; Weedon et al. 2017). The novelty of fake news and the negative feelings they often cause are additional reasons that fake news travels faster than real news (Vosoughi et al.

2018). At the same time, the decentralized distribution makes it difficult to stop the dissemination of fake news since there is no longer a central origin of spread that can be addressed. Fourth, there is no authority that controls the authenticity of information (Shao et al. 2016). The process of journalistic information verification and the associated gatekeeping functions can no longer be guaranteed because social media allows everyone to participate in the production of news (Allcott and Gentzkow 2017; Rubin et al. 2015). Due to the lack of control, the responsibility for validating the authenticity of information now lies with the users (Rubin et al. 2015).

In recent times, research has increasingly focused on why users can or cannot detect false information, especially in social media. A key factor for validating authenticity is the credibility of information, which is complex and diverse (Wathen and Burkell 2002; Petty and Cacioppo 1986). Source (expertise, knowledge, and trustworthiness), receiver (receptivity for a message), and message characteristics (consistency and clarity) all interact in the assessment of the credibility of information (Wathen and Burkell 2002). Additionally, the believes and prior knowledge (Pennycook et al. 2018) of a user are important because these factors can lead to confirmation bias. People prefer to believe information that matches their preexisting beliefs (Allcott and Gentzkow 2017; Kim and Dennis 2019; Housholder and LaMarre 2014). When they encounter information that challenges their preexisting beliefs, they experience cognitive dissonance (Mills 2019; Festinger 1957). Resolving this cognitive dissonance takes cognitive effort, and people tend to avoid this (Simon 1979). This tendency to favor information that confirms one's preexisting beliefs and ignore information that challenges them is called confirmation bias (Devine et al. 1990; Koriat et al. 1980; Nickerson 1998). At the same time, the assessment of other social media participants also has an impact on the perception of the credibility of news articles (Kim et al. 2019). Despite these factors, the ability of people to decide whether the information is true or false has some weaknesses naturally, especially in the social media environment (Shu et al. 2017; Lazer et al. 2018; Moravec et al. 2019). It is, therefore, necessary to develop countermeasures that support social media users in handling fake news.

4.2. Improving Users’ Fake News Recognition in Social Media – an Online Experiment

No technical solution has been able to sufficiently detect fake news in social media. Thus, an alternative approach is required, which may influence the credibility perception among the audience itself. Fact-checking, source credibility ratings, and the specific design and format of news sources have all been identified as possible means of fake news detection (Amin et al. 2009; Kim and Dennis 2019). However, previous research has raised concerns about the effectiveness of these approaches (Lazer et al. 2018; Pennycook et al. 2018). Assessing the source has been found to be an important and helpful factor for deciding whether a news post is true or false, with known sources often creating more trust in the presented news (Amin et al. 2009; Kim and Dennis 2019). However, anyone can publish news in social media, and therefore large amounts of content are published by unknown sources. Consequently, it is all the more important to analyze possibilities to improve fake news detection in situations where the source is unknown (Kaplan and Haenlein 2010). Further, Lazer et al. (2018) address the necessity of multidisciplinary approaches when attempting to detect fake news. These include structural changes as well as empowering individuals.

To empower individuals in a way that supports fake news detection, one needs to understand how human perception, processing, and sense-making of information and content work. One of the influential factors in the perception process is knowledge (Goldstein 2010), which enables information perception and thus plays a key role in whether one classifies information as true or false.

Research has also shown that the increase in a person’s prior knowledge leads to improved sense-making in news consumption and better handling of large amounts of information (Pentina and Tarafdar 2014). In the specific context of fake news detection, a person’s knowledge base has shown to be one of the core resources used for assessing the truthfulness of news (Flintham et al. 2018). Hence, improving a user’s knowledge base on a given topic in order to increase analytical thinking in social media holds promising potential in the fight against fake news (Bronstein et al. 2019).

When assessing possibilities to reach an enhancement in user behavior, there is currently much discussion about the approach of digital nudging (Thaler and Sunstein 2008; Weinmann et al. 2016). As Sunstein observes, the nudge *disclosure* can improve a user’s knowledge. In social media, such a digital nudge could be implemented in the form of related articles, which address the same subject as the main article but do not necessarily take the same point of view (Sunstein 2014). This may help provide users with more information on a particular topic and, hence, empowers them to better detect fake news through knowledge improvement.

Therefore, the purpose of our research is to investigate whether related articles improve fake news detection by users in social media environments. We investigate three possibilities on how to provide related articles by conducting an online experiment simulating the Facebook news feed. Our results suggest that related articles have a significant effect on improved fake news detection. We achieved the best results when providing a mix of controversial related articles.

First, we provide the theoretical background on news consumption, digital nudging, and the derivation of the research hypotheses. Next, we expand on the experimental method used to address our hypotheses. Subsequently, we present the results of the experiment and conclude by discussing results, highlighting the theoretical and practical implications, and pointing out the limitations of our research.

Theoretical Background and Hypotheses Development

News Consumption and Fake News Detection in Social Media

It is important to understand how users consume news in social media to improve a user's ability to detect fake news. According to Pentina and Tarafdar (2014), the way news is consumed via social media has changed drastically compared to consumption via more conventional media. They theorize that sense-making of news in social media environments and the consequential formation of knowledge relies on information overload strategies. Two specific aspects characterize news consumption (Pentina and Tarafdar 2014):

First, the screening news stimuli explains which channels, sources, and content are considered by an individual. Second, sense-making relies on processing and interpreting information from news, which involves the interpretation of the meaning of the obtained news and its transformation into knowledge.

Social media is a news provider, which offers a diversity of opinion and social legitimacy, but simultaneously, it also offers space for subjective opinions and fake news stories (Pentina and Tarafdar 2014). This poses a severe challenge due to the fact that humans are fairly ineffective at recognizing deception (Conroy et al. 2015). This ineffectiveness is caused by humans' frequent inability to determine whether the information presented to them is true or false (Shu et al. 2017). Fake news detection is the correct decision of an individual that information is false. However, in the context of decision-making, natural human deficits resulting from cognitive and behavioral biases often lead to erroneous assessments (Hansen and Jespersen 2013).

Throughout the cognitive processing and interpreting of news information, news consumers evaluate trustworthiness and reliability (Pentina and Tarafdar 2014). Therefore, individuals access a variety of sense-making strategies, like source reliability or comparing news content to their own knowledge to evaluate the credibility of news (Flintham et al. 2018).

However, prior exposure to false knowledge related to a news statement also increases the believability of fake news. That is, due to the *illusory truth effect*, individuals tend to misinterpret fake news as true when they were exposed to the false knowledge beforehand (Pennycook et al. 2018). From a positive point of view, it also indicates that improving users' knowledge while consuming news in social media environments improves their ability to decide which news is true or fake correctly. In contrast to fake news, truthful news is those stories that cover content that is verifiable through data, facts, and research. Even though the notion of *true news stories* is widely used, one still has to consider that journalism of this kind will always include some extent of a personal footprint, such as through the author's selection of topics covered (Hermida 2012).

Related Articles

Researchers have shown growing interest in ways to influence decision-making actively in situations strongly affected by biases. This type of influencing behavior is often referred to as soft or asymmetric paternalism or nudging (Leonard 2008). Thaler and Sunstein define a nudge as "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives." (Thaler and Sunstein 2008, p. 6). Nudges target biases and attempt to overcome them by conscientiously altering the *choice architecture* (design of possible choices) presented to the respective target group.

The goal is to foster decision-making contexts that promote behavior that is beneficial both to individuals and to society (Thaler and Sunstein 2003). Nudges should not hinder freedom of choice and should aim to make life simpler, safer, or easier for people to navigate while remaining transparent and open rather than hidden (Sunstein 2014). Where the nudging philosophy has been transferred into the digital environment, it is referred to as *digital nudging* (Weinmann et al. 2016). Common environments for digital nudging are social media platforms (Weinmann et al. 2016) such as Facebook, Twitter, or Instagram. Social media providers are choice architects who choose the content and format of the information presented to their users. Nudges could, therefore, be used to improve fake news detection in social media environments and become a countermeasure against fake news.

Sunstein presents a list of ten important *nudges*, one of which is disclosure (Sunstein 2014). The disclosure describes the adding of supplementary information to a specific topic or situation, such as nutritional details on the packaging of food. In the context of digital disclosure, it is explained that “more detailed and fuller disclosure might be made available online for those who are interested in it” (Sunstein 2014, p. 586). Yet, simplicity remains essential, which means that information presented must both be comprehensible and accessible.

Since 2017, Facebook has been working with a feature called related articles, which shows users a range of articles under an original post to help them make decisions that are more informed. This feature is mainly used in the U.S., although there have been plans for a future rollout in Germany, France, and the Netherlands (Chaykowski 2017). From a theoretical perspective, related articles are a type of disclosure nudge.

Disclosure has already been used successfully in the context of social media, e.g., in the field of privacy protection (Yang et al. 2013). We hypothesize that disclosure is a promising nudging approach to improve a user’s knowledge, and can consequently stimulate analytical thinking. This is because users have access to more information and can better reflect the news, leading to an improvement in fake news detection (Bronstein et al. 2019).

News can be either true or false (actual state) when looking at news consumption in social media. For both situations, the user can further perceive the news post as true or false (perceived state). The results of this classification task can be described with the help of specificity. Specificity describes the proportion of actual fake news that is detected to be false by the users. We summarize this within our first hypothesis.

H 4.2.1 The presence of related articles improves fake news detection in terms of improved specificity.

Design of Related Articles

There are various ways to implement a disclosure nudge using related articles. Since fake news is published to spread false information, the easiest way to recognize them is by knowing the truth (Shu et al. 2017). Hence, when knowing all facts on a given topic, detecting fake news is simple.

From this point of view, users may better detect fake news as such, while articles telling the truth are available to them. Following a modified form of the *illusory truth effect* (Pennycook et al. 2018), exposure to true statements may enable users to detect fake news even though they do not know the truth. Therefore, a disclosure nudge in the form of related articles that solely tell the truth ought to increase fake news detection.

Thus, we assume that related articles that are all truly achieve the best results in fake news detection, measured by the highest specificity.

H 4.2.2 Related articles that tell the truth lead to an increase in specificity in fake news detection compared to having no related articles.

Finding verifiably true related articles is a costly and challenging task for the platform provider, and it likely is impossible to automate fully (Shu et al. 2017).

An alternative is the provision of controversial news, which is easier to implement. Current technical approaches exist that build on detecting controversies in the news (Vicario et al. 2019). Controversial news in social media represents those topics that find supporters for conflicting sides of an argument and are debated heatedly (Garimella et al. 2018). A popular example of such a controversy would be believing and disbelieving that climate change is real. From a user's perspective, previous studies have already demonstrated that controversial discussions with supporting and opposing arguments (e.g., pro and contra the existence of climate change) help to improve people's analytical thinking. Results suggest that people presented with point/counterpoint arguments are less susceptible to biases (Wiley 2005). Controversial news, in addition, encourage users to actively think about a certain topic, which has previously been shown by analyzing online search behavior (Weeks et al. 2012). Thus, controversial related articles may increase active, open-minded, and analytical thinking, which is expected to be promising in the fight against fake news (Bronstein et al. 2019).

Controversy can be implemented in different ways when designing related articles. Firstly, in the sense of controversial related articles that show different viewpoints on a given topic.

Herein, the collection of related articles itself includes controversy, with some articles agreeing and some disagreeing with the main article. Research has shown the potential of offering balanced viewpoints on a topic to counter cognitive biases and thus hinting at the potential of mixed, controversial related articles in the fight against fake news (Caraban et al. 2019). Hence, we hypothesize that a mix of controversial related articles also increases fake news detection in terms of higher specificity compared to users who do not receive controversial related articles.

H 4.2.3 A mix of controversial related articles leads to an increase in specificity in fake news detection compared to having no related articles.

Secondly, controversy can be achieved by using related articles that oppose the main article in the strongest possible way. ‘Considering the opposite’ is a strategy to overcome cognitive biases that has been thoroughly discussed in research in psychology (Lord et al. 1984). The strategy consists of directly pointing people to the opposite perspective on a given topic or question. Consequentially, we hypothesize that designing controversial related articles in a way that they strongly oppose the main article and thus hint at the opposite alternative increases fake news detection through an increase in specificity as compared to users who do not receive controversial related articles.

H 4.2.4 Opposing related articles lead to an increase in specificity in fake news detection compared to having no related articles.

Note that H 4.2.2, H 4.2.3, and H 4.2.4 suggest that three mutually exclusive versions of related articles are all effective. There is no line of theoretical argumentation unequivocally suggesting which of the three designs of related articles performs best. In case multiple hypotheses are empirically supported, it is an empirical matter to identify which of the designs performs best.

Besides improving specificity, it is important not to aggravate sensitivity. Sensitivity describes the proportion of actual true news that is detected to be true news by the user. We do not hypothesize any effect, but we include sensitivity in our analysis and test for possible aggravation.

Experiment Design and Procedures

To test our hypotheses, we conducted an online experiment that assesses the impact of related articles on fake news detection. The experiment included presenting an interactive news feed similar to that of Facebook. The participants could scroll through the news feed and interact to achieve a scenario that is as realistic as possible. Facebook was used as a template because it is the most used social media platform and therefore provides a real-life, natural, and known setting (Newman et al. 2018). All parts of the experiment were in the German language.

The experiment consisted of six stages: (1) introduction, (2) questions on demographic factors, (3) introductory news feed, (4) natural interaction run, (5) questions on fake news detection, and (6) debriefing. Appendix B.1 summarizes the experiment procedure.

Throughout the first four stages, the context of fake news was not explained to avoid biases in the participants' behavior. Rather, the context mentioned was that of general online behavior in social media. All participants were asked demographic questions to help us develop a clear picture of who completed the experiment. Our questions focused on age, gender, education, current profession, and intensity of social media usage. In order to make sure all participants were accustomed to the functionality of the feed, an introductory news feed was presented using a single news post, in which all functions (like, share, comment, report, open articles) had to be tested.

Afterward, all participants went through a natural interaction run with the actual news feed, which consisted of news posts such as those published by newspapers or news services. The participants received six news posts about current events (politics, environment, and celebrities), of which three were truthful news, and three were fake news. Each of the three topic categories was covered via one truthful and one fake news post in order to avoid a bias caused by topic selection. Further, all six articles are real archived news articles that were published online. For each article - and also all related articles - we have conducted a review of fact-checking websites (e.g., snopes.org) to ensure that it is either true or false news. As the news source strongly impacts the perception of credibility, we chose six articles from rather unknown sources (Kim and Dennis 2019). In addition, we also avoided to include other criteria that would allow participants to clearly identify a fake news article as such (e.g., spelling mistakes). We did this to minimize other external effects in order to better observe the change in specificity caused by related articles. The order of the posts varied, but all participants were presented with the six identical news stories. Appendix B.2 gives an overview of all main and related article headlines.

All participants were randomly assigned to one of four treatments:

- **Treatment 1 (T1) - Control Group**: Six main articles are shown. No related articles are shown.
- **Treatment 2 (T2) - True**: Six main articles are shown. Four true related articles are shown under each main article, regardless of whether the main article is true or fake.
- **Treatment 3 (T3) - Mixed**: Six main articles are shown. Two fake and two true related articles are shown under each main article, regardless of whether the main article is true or fake.
- **Treatment 4 (T4) - Opposing**: Six main articles are shown. For each main article, four articles with opposing content (in relation to the main article) are shown. Consequently, four true related articles are shown for each fake main post, and four fake related articles are shown for each true main post.

Our control group T1 allows for comparing the general effect of related articles vs. no related articles (H 4.2.1). The treatment groups T2 to T4 are used to assess H 4.2.2, H 4.2.3, and H 4.2.4. T2 includes the highest degree of truthfulness in the related articles, with all related articles showing the truth (H 4.2.2). While T3 offers the highest controversy within the related articles (two true, two fake – H 4.2.3), T4 portrays the largest possible controversy between the main article and the related articles (H 4.2.4). One must consider that in the case of a fake main article, the related articles in T2 and T4 are identical (four true related articles in both cases). For all four treatments, the news feed allowed us to track the activities (liking, sharing, commenting, reporting, opening articles) executed during the experiment, which enabled us to evaluate the social media usage behavior of the participants.

Figure 9 shows an example of a real news post, including related articles. Clicking on either the main or the related article opened the full text of the article.

After the news feed interaction, all participants were shown the original six news posts (without related articles) and were asked to explicitly state for each post whether they regard it as a fake or true article. This step provided the basis for evaluating the individual fake news detection abilities.

Figure 9: True news post including related articles

Finally, the experiment ended with a debriefing session. We informed the participants about the fact that we had manipulated different features of the posts, which not necessarily correspond to the real world.

Participants for the experiment were recruited via multiple channels. We used Facebook as the main source for recruiting and linked to our experiment in various groups, mainly targeting students and young professionals. Among others, the targeted groups included sports groups, student associations, university groups, and the authors' social network. Other sources of participants were messenger providers such as WhatsApp or email lists in a university and work context. The choice of recruiting channels was selected because the majority of Facebook users in Germany are aged between 25 and 34 (Statista 2018b). Thus, we ensured that participants are used to social media environments and are, in fact, part of the target audience mainly affected by fake news in social media.

Results

Characteristics of the Sample

In total, 322 people completed our experiment, with 146 female and 176 male participants. Thereof, 311 participants are between 18 and 35 years old, and 11 are older than 35. Regarding education, 221 participants have at least one academic degree. Analyzing the participants' use of social media, we found that 44.7% use social media daily, and 92.9% use Facebook. Consequently, we conclude that participants are familiar with the use of social media, such as Facebook.

Seventy-seven participants were randomly assigned to treatment T1 by the assignment algorithm of the survey provider, 77 to T2, 93 to T3, and 75 to T4. To ensure comparability between treatments, we analyze their homogeneity. For discrete variables, we use chi-squared-tests and analysis of variance tests (ANOVA) for continuous variables. The test statistics show that participants do not differ significantly between the treatments in terms of gender, age, social media usage, or education. Therefore, the groups can be assumed to be comparable.

68.0% of the participants actively interacted with the news feed, meaning they either shared, commented, liked, opened, or reported articles. A range of 1 to 34 interactions per active user was recorded, with a mean of 3.4 and a median of 2.0. Previous studies of social media usage report a 52% activity rate among users (Statista 2018a). Based on this figure, participants in our study were slightly more active than the average user.

Participants shared between 0 and 6 posts, with a mean of 0.2 and a median of 0. They commented on 0 to 6 posts, with a mean of 0.7 and a median of 0. They liked 0 to 6 posts, with a mean of 0.4 and a median of 0. Participants who had the related articles feature implemented (T2-T4) opened related articles between 0 and 19 times, with a mean of 0.8 and a median of 0. On average, participants in T2, T3, and T4 used the news feed for a longer period of time. Participants assigned to treatment T1 interacted with the news feed for a mean of 80 seconds and a median of 62 seconds, T2 for a mean of 110 seconds and a median of 66 seconds, T3 for a mean of 92 seconds and a median of 77 seconds and T4 for a mean of 109 seconds and a median of 77 seconds.

Overall, participants reported an average of 0.7 articles with a median of 0 and a spread from 0 to 6. 64.6% of the participants did not report a single article.

For all interactions mentioned, we test whether there are significant differences between the treatments. Interactions (like, share, comment, report, or open related articles) do not differ significantly between treatments, excluding T1 for related articles as they were not available to this treatment.

Fake News Detection Behavior

After interacting with the news feed, in stage 5, all participants were asked to explicitly state whether they considered each of the six posts to contain fake or true news. The above-mentioned task represents a classification task consisting of the true state of the article (true or fake) and the participant's perception of the article (true or fake). As an adequate and common tool for comparing different approaches to classification, the confusion matrices, including true positives (TP), false negatives (FN), false positives (FP) and true negatives (TN) of all four treatments are constructed (Sokolova and Lapalme 2009). Table 7 shows an exemplary confusion matrix for participants in T1. The overall n of that matrix describes the number of participants in T1 * 6 (every participant was shown six news posts for classification).

	Perceived True	Perceived Fake
Actual True	182 (TP)	49 (FN)
Actual Fake	37 (FP)	194 (TN)

Table 7: Confusion matrix of T1

The same information can be extracted for all treatments, considering six classifications for each participant in each treatment group. To generally assess the difference between showing and not showing related articles, the confusion matrices of T2-T4 can be aggregated. Table 8 shows the confusion matrix of all groups in an aggregated manner.

	T1: Control	T2: True	T3: Mixed	T4: Opposing
TP	182	169	218	177
FN	49	56	61	54
FP	37	26	25	36
TN	194	199	254	195

Table 8: Classification results of all groups

From these responses, specificity ($\frac{TN}{TN+FP}$) and sensitivity ($\frac{TP}{TP+FN}$) values are calculated for all four groups as a performance indicator. The respective results are presented in Table 9. For both specificity and sensitivity, higher values are preferable, as they indicate better classification results.

Observing the performance results, Table 9 shows an increase in specificity, when comparing the aggregated results of all groups being shown related articles (T2-T4) with T1. Looking at differences between individual treatments, especially between T1 as compared to T2 and T3, T3 (mixed) exhibits the highest specificity, followed by T2. T4 shows the smallest increase in specificity when comparing the different alternatives for designing related articles. Between all groups that show related articles in the experiment, T3, with its high levels of controversy among related articles, performs best with respect to both specificity and sensitivity.

	Specificity	Sensitivity
T1: Control	83.98%	78.79%
T2: True	88.44%	75.11%
T3: Mixed	91.04%	78.14%
T4: Opposing	84.42%	76.62%
T2-T4 (aggregated)	88.16%	76.73%

Table 9: Performance metrics

In the next step, the significance of the observed performance differences is assessed. The tests are performed to evaluate whether the increase in specificity and sensitivity are likely caused by the treatment. As three fake and three true articles were shown to each participant, there exist only four levels for both accuracy metrics (0 to 3 out of 3 articles classified correctly) for each individual participant. We, therefore, use chi-squared-tests to determine whether significant differences exist between the treatment groups.

As H 4.2.1 regards the general effectiveness of related articles, we perform a chi-squared-test on the specificity and sensitivity values between T1 and T2-T4. While there are no significant differences in treatment sensitivity (p=0.632), specificity is significantly lower (significance level 10%) in the control group as compared to treatments T2-T4 (p=0.063). Thus, we find support for H 4.2.1. There is a positive effect, as people being shown related articles perform significantly better at detecting fake news. Simultaneously, the decrease in sensitivity observed from no related articles to related articles (78.79% to 76.73%) is not significant, so we do not find an effect of related articles deteriorating one's ability to detect true news posts as such.

To evaluate the impact of the degree of truthfulness and controversy in the news posts, we conduct pairwise comparisons between all treatment groups using chi-squared tests. The aim is to test H 4.2.2, H 4.2.3, and H 4.2.4. Table 10 shows the resulting p-values of all pairwise comparisons for specificity and sensitivity. The upper-right triangle in Table 10 shows the results for specificity, the lower-left triangle for sensitivity.

	T1: Control	*T2: True*	*T3: Mixed*	*T4: Opposing*
T1	-	0.178	0.008	0.178
T2	0.543	-	0.449	0.306
T3	0.934	0.748	-	0.029
T4	0.571	0.939	0.885	-

Table 10: P-values of pairwise comparisons for specificity (upper-right triangle) and sensitivity (lower-left triangle, grey)

Looking at all significant values (significance level 5%) in Table 10, we conclude that participants in T2 did not perform significantly better than participants in the control group (T1). Therefore, we can reject H 4.2.2. Truly related articles do not necessarily lead to improved fake news detection.

Beyond that, the results indicate that specificity levels are significantly higher in T3 as compared to T1 and T4. Participants in T3 achieved better results in detecting and classifying fake news. This supports H 4.2.3 and, thus, the positive influence of controversy within the related articles when trying to improve fake news detection abilities.

Furthermore, the slight increase in specificity from T1 to T4 appears not to be caused by the treatment, as the corresponding p-value is very high. Consequently, we can reject H 4.2.4 and the assumption that related articles, which strongly contradict the main article, lead to an improvement in users' fake news detection abilities.

Sensitivity is not affected, as all pairwise comparisons yield high p-values. This indicates that while related articles improve fake news detection behavior under specific conditions, they seem not to deteriorate the subjects' ability to classify true news as such. The results presented above are also robust when equalizing the treatment group sizes (T3 n=76) through random sampling and then conducting the analyses as before.

Discussion

Contributions

This chapter details an online experiment to test the influence of related articles on fake news detection in social media. The results indicate that related articles, in general, can improve fake news detection. More precisely, a mixture of controversial articles under the main article fosters the ability to detect and classify fake news articles.

In response to our first hypothesis (H 4.2.1), this study finds an indication that related articles have a positive impact on fake news detection behavior in the context of social media. Related articles portray a digital nudge, and fall under the category of the disclosure. The key characteristic of this nudge is that it supplies additional knowledge and triggers analytical thinking to those who read the news post (Sunstein 2014). Disclosure-type nudges have proven to be effective at improving peoples' behavior in several environments (Ho 2012; Kroese et al. 2016). Our study demonstrates that this positive effect also applies in the context of fake news detection.

However, in response to hypothesis H 4.2.2, our results indicate that providing solely true related articles does not necessarily enable better fake news detection. Thus, we cannot confirm our assumption of exposure of truthful knowledge (see Pennycook et al. 2018) being an effective approach. Also, showing related articles that strongly contradict the main article (H 4.2.4) does not lead to a significant improvement in fake news detection.

In any case, automatically providing only related articles that are true seems almost impossible in practice (Shu et al. 2017). For this reason, we hypothesized that a mix of controversial related articles might be an effective approach (H 4.2.3). Our results indicate that a mix of controversial related articles enables users to classify better and detect fake news articles as they achieved significantly higher specificity than treatments T1 and T4. This is in line with previous studies, which demonstrated that people presented with point/counterpoint arguments are less susceptible to biases (Wiley 2005). Thus, a mix of controversy may lead users to reflect more thoroughly on information, rather than simply accepting the information as it is presented in the main article.

Moreover, controversial articles posted below true articles do not reduce users' ability to recognize and classify true articles, which is immensely important if the nudge is to have the desired effect.

In addition to these theoretical contributions, the results of our study also have practical implications. For social media platform operators, our results suggest a simple means of supporting their users in fake news detection. In particular, we provide valuable and specific insights as to how the feature might be implemented, as the use of only true articles in this type of feature would be complex and costly, if not impossible, to achieve (Conroy et al. 2015). In contrast, the use of a mix of controversial articles (as in the mixed treatment T3) is a much more practical and feasible approach. Based on the results of our study, this is the most effective way to improve the accuracy of fake news detection. Further, the usage of related articles increased users' screen time in our study and, thus, can be deemed compatible with common social media platform business models.

Overall, this chapter demonstrates empirically that a specific form of related articles improves fake news detection in social media, while not compromising the users' ability to identify truthful news as such.

Limitations and Further Research

Our study has some limitations, which highlight the need for future research. First, the generalizability of our research needs to be validated in future research. This is because the age range of our participants was limited to relatively young Germans. Further tests with different age groups and nationalities should be carried out. However, since our survey targets social media, the average young age of participants fits the user group, as a recent study showed that the largest share of Facebook users in Germany is aged between 25 and 34 years old (Statista 2018b).

Further, there are various factors crucial to the detection of fake news (Flintham et al. 2018). Prior research has shown that the source of an article is an important factor for its perceived credibility (Amin et al. 2009). Even though the impact of source credibility is purposely omitted from the scope of this study, it should be further investigated in future research whether the observed effects hold even when the news come from known sources.

The participants' previous knowledge of the topics covered in the articles may have influenced detection behavior, and should also be a focus of future research. However, with random assignment to treatments, this can hardly explain the observed treatment effects. In addition to that, not all selected related articles are of the exact same relevance to the main article. The choice of related articles in T3, therefore, potentially impacts the perceived helpfulness of this treatment.

Our study focuses on short-term behavior. Future attempts should investigate long-term behavior in order to provide a more holistic understanding of behavioral patterns. In this sense, it will also be important to investigate users' reactions to the related articles disclosure nudge once it has been in use for an extended period of time and is no longer a novel feature.

Lastly, although we have attempted to authentically recreate the original news feed, our study took place in an experimental environment, a fact that may have influenced the behavior of our participants. It would be insightful to gather data on real social media usage to address this issue.

Overall, and despite these limitations, our study will help with creating a better environment in terms of improved detection of fake news.

Nonetheless, research needs to investigate the impact of other digital nudges and to evaluate their effectiveness, not only in terms of improving fake news detection but also in terms of reporting suspicious content. There are promising approaches (such as SNs) in this field, and one might potentially combine multiple nudges to combat fake news more effectively.

To conclude, our study indicates that a mix of controversial related articles improves users' ability to detect fake news. Related articles are a comparatively simple concept and are already present on some platforms. They may prove to be a valuable weapon in the fight against fake news. Our finding that a mix of controversial articles is effective in this context makes the use of related articles much more feasible, as this removes the need to ensure that all related articles are true or opposing. The use of digital nudges may significantly advance the fight against fake news, and ideally, these will be combined with the development of automated, IT-centered solutions. In the long run, the problem of fake news will not be easy to solve, particularly with respect to political manipulation, which poses a central challenge for society. Consequently, a variety of approaches will be needed to prevent the spread of false and misleading information via social media in the future.

4.3. Improving Users' Fake News Reporting in Social Media – two Online Experiments

IS research has proven that enabling social media users to recognize false and misleading information better is a promising approach (e.g., Kim et al. 2019 or Chapter 4.2 of this thesis). However, countermeasures like source-ratings by experts (Kim et al. 2019) and even stricter approaches such as deleting accounts that spread mainly fake news require a pre-check of suspicious content. Since the detection of fake news is difficult or impossible to automate (Chatfield et al. 2017; Shu et al. 2017), human effort is required. As the content on social media platforms is crowdsourced, it is compelling to crowdsource the pre-selection of suspicious content without financial incentives. Even if user ratings and reports might be subject to biases (Kim et al. 2019), online social network services rely heavily on their users' assessment of trustworthiness (Frenkel and Maheshwari 2018) and their reporting of fake news for the pre-selection process (Mosseri 2016; Jamieson and Solon 2016). It is, therefore, essential to guide the behavior of users in such a way that they actively and accurately report fake news and thus inform social media platform providers about potentially inauthentic content.

The social environment of an individual heavily influences behavior. Social psychologists have shown that SNs are an important factor in the decision-making process (Deutsch and Gerard 1955; Cialdini et al. 1990). By changing the perception of a SN, the behavior of an individual can be guided. That is, SNs can be used to modify a human's environment to encourage prosocial behavior (Cialdini et al. 1990; Cialdini 2003; Cialdini et al. 2006). There are two types of SNs: a injunctive SN describes which behavior most people approve or disapprove. A descriptive SN describes what other individuals do in particular situations (Cialdini et al. 1990). The more individuals demonstrate a certain behavior, the higher is the strength of a descriptive SN (Cialdini et al. 1991).

In this chapter, we examine the effect of injunctive and descriptive SNs on social media users' reporting of fake news. Our motivation lies in the nexus of (1) the importance of fighting fake news (Bernard et al. 2019; Kim et al. 2019; Moravec et al. 2019), (2) the general power of SNs to steer behavior in a direction that is beneficial for individuals and society (Cialdini et al. 1990; Cialdini et al. 1991; Demarque et al. 2015; Kormos et al. 2014), (3) the role of user reporting in other domains of online activities such as online abuse and cyberbullying (Wong et al. 2016), and (4) the fact that the positive and potential negative effects of applying SNs to fake news reporting in social media have not been studied yet. Knowing that SNs improve fake news reporting, they would allow for a rather low cost and effective tool in the overall approach to fight fake news. On this ground, we derive theoretical hypotheses and conduct two online experiments on the reporting behavior of social media users in the presence of SNs and fake news. With this, we follow a recent call for more IS research on socio-technical interventions as remedies against fake news (Bernard et al. 2019). Based on a combination of theoretical arguments and empirical evidence, we suggest that especially the combination of injunctive and descriptive SN messages as social media design features has positive effects on fake news reporting.

Theoretical Background and Hypotheses Development

Countermeasures against Fake News

Many social media platform providers are criticized for the lack of countermeasures against fake news (The Editorial Board 2016). For example, Twitter published a website for the United States midterm elections 2018, which already contained fake news stories shortly after its launch (Warzel and Mac 2018). Therefore, social media providers tried to develop various countermeasures to prevent the spread of fake news, such as automatically detecting social bots or cyborg users. However, current approaches do not automatically check the authenticity of information (Weedon et al. 2017). Due to their complex structure, the detection of fake news is difficult to automate (Chatfield et al. 2017; Shu et al. 2017; Lazer et al. 2018). Therefore, social media providers work with third-party fact-checking organizations who manually check suspicious information (Hunt 2017). The social media provider then can flag the corresponding content and display other articles on this topic. These countermeasures should support users by bringing fake news directly to their attention and providing more diversified information. Another method is to reduce the financial incentives to spread fake news by prohibiting sites from advertising if they have repeatedly shared fake news (Shukla and Lyons 2017). All of these approaches depend on a pre-selection of suspicious content. Therefore, social media providers rely on the support of their users (Murray 2016). Facebook, for example, passes content to third-party fact-checking organizations for verification, if enough users report a specific content as fake (Mosseri 2016). The users and their reporting behavior become a central element of the system because they act as enablers for the subsequent mechanisms. However, no data on the reporting behavior of users, especially in the context of fake news, is available or publicly accessible (Mustafaraj and Metaxas 2017).

From a theoretical perspective, previous research on the bystander effect shows that people's willingness to help is reduced the more other people are present (Fischer et al. 2011). This effect applies not only in the case of physical presence (Latan and Nida 1981; Peter et al. 1972) but also in an online environment (Fischer et al. 2011). Following the bystander intervention model, a bystander must follow five steps, which are noticing and appraising an event as an emergency, recognizing the own responsibility, deciding to intervene, and finally intervene (Latané and Darley 1970). These steps can be affected by the presence of other people. For instance, a larger group size leads to reduced response rates and willingness to help in online communities (Voelpel et al. 2008) or during e-mail communication (Barron and Yechiam 2002; Blair et al. 2005). In the context of social media, where many users see the same content, this could lead to a diffusion of responsibility (Voelpel et al. 2008), where the individual does not take responsibility for reporting fake news. The missing feeling for accountability on the user's side is another factor, which increases this effect. Within a group of people, an individual evaluates its own accountability lower because another individual may be more accountable to intervene in a specific situation. Making accountability transparent may reduce this effect. However, transparency can also have a negative effect when users feel that their intervention is evaluated by others and, therefore, do not dare to take action (DiFranzo et al. 2018). For these reasons and because of the importance for other countermeasures, it is necessary to show to social media users that reporting is a desirable and appropriate behavior in order to increase reporting of fake news in social media.

Social Norms

In the past, research has already proven the influence of social environments on an individual's behavior. According to social psychologists, so-called SNs are an appropriate tool to foster prosocial human behavior (for instance, Cialdini et al. 1990; Cialdini 2003; Goldstein et al. 2008). This is because people regularly use the perception and judgment of other people as a reliable source of evidence (Deutsch and Gerard 1955). A special characteristic of social media is that decision-making is not isolated but takes place in a virtual social environment with SNs. Within such a social environment, the attitude and behavior of others can have a strong influence on an individual's behavior (Cialdini et al. 1990; Deutsch and Gerard 1955) and may lead to prosocial activities (Kuem et al. 2017).

SNs are "an individual's beliefs about the typical and condoned behavior in a given situation" (Kormos et al. 2014, p. 480). Research distinguishes two types of SNs. First, injunctive SNs "specify what ought to be done" (Cialdini et al. 1990, p. 1015) and describe the behavior, which is expected from the social environment by defining which actions are desired and which are undesired (Reno et al. 1993). Second, descriptive SNs inform about what others typically do in a particular situation (Cialdini et al. 1990). Although both types of SNs provide social information, they act through different sources of motivation (Cialdini et al. 1990; Reno et al. 1993; Rimal and Real 2003). Injunctive SNs influence behavior by highlighting potential social rewards and punishments for desirable or undesirable behavior (Cialdini et al. 1991). Such a moral evaluation influences the behavior even if the imaginary others are not family or friends, but are attributed to the general construct of society member (Cialdini 2007).

Although SNs mostly guide an individual's behavior, they do not always and in every situation unfold the desired effect (Cialdini et al. 1990). The focus theory of normative conduct (Cialdini et al. 1990) assumes that an individual is more likely to behave according to a SN when the individual's attention is focused on the norm message at the moment of behavior. Focusing attention on a SN is required to activate the SN and, thus, increases the likelihood of desired behavior. Priming or highlighting the behavior of other individuals are exemplary tools to do so (Cialdini et al. 1990; Goldstein et al. 2008).

Previous research already demonstrated the importance of injunctive SNs for decision-making in different contexts such as alcohol use (LaBrie et al. 2010; Rimal and Real 2003), gambling (Neighbors et al. 2007), tax evasion (Wenzel 2004), organ donation (Park and Smith 2007) and risky sexual online behavior (Baumgartner et al. 2011). Due to these versatile influences on behavior, we suspect that injunctive SNs can also be used to improve the reporting behavior of social media users. We expect that when social media users are made aware that reporting fake news is a behavior desired by the social environment, it also positively affects their decision to report fake news. In particular, we expect that the use of an injunctive SN will lead to more reported fake news stories. Based on this, we derive the following hypothesis:

H 4.3.1 The presence of an injunctive SN message indicating that reporting fake news is a socially desirable behavior increases the amount of reported fake news.

Contrary to injunctive SNs, descriptive SNs motivate actions by providing information about the behavior of others and therefore showing which actions are likely to be effective, adaptive, and appropriate (Cialdini et al. 1991). The behavior of others is used as a heuristic, giving the decision-maker an information-processing advantage and decisional shortcut in situations that are uncertain, ambiguous, or threatening (Jacobson et al. 2011). Previous research has demonstrated the successful application of descriptive SNs to reduce littering (Cialdini et al. 1990; Reno et al. 1993), support recycling (Nigbur et al. 2010), improve voting behavior (Gerber and Rogers 2009), promote towel reuse (Goldstein et al. 2008), increase monetary donations (Cialdini 2007; Croson et al. 2009), and to encourage healthier eating (Robinson et al. 2014). Based on these results, we suspect that the demonstration of others reporting contents on social media platforms convinces the user that reporting is appropriate behavior. Therefore, we suspect that the number of reports is increasing compared to the presence of no SNs at all. This assumption leads to the following hypothesis:

H 4.3.2 The presence of a descriptive SN indicating that other people report fake news on social media increases the amount of reported fake news.

Bringing the injunctive SN to attention by means of a descriptive SN can further improve behavior (Cialdini et al. 1990). Following focus theory, it suggests that the presence of a descriptive norm can (re-)direct the user's focus to the injunctive norm. In a study on recycling behavior, Cialdini (Cialdini 2003) demonstrates that the joint use of injunctive and descriptive SNs leads to a particularly high intention to recycle. Similarly, Schultz et al. (Schultz et al. 2008) reveal in a field experiment on towel reuse that a combination of the two types of SNs has a stronger impact on behavior than just using one of them. The results of their study indicate that a combination of descriptive and injunctive SNs not only activates an existing norm, but it is also used as a guide for one’s behavior. Applying these insights to the case of fake news reporting, we expect that the presence of a descriptive SN – describing other users having reported news posts – will focus a user’s attention to the injunctive SN – describing that reporting is the desired behavior – and, thus, the amount of reported fake news will lead to a greater increase in the number of fake news stories and thus to an overall increase in the number of reports and formulate the following hypothesis:

H 4.3.3 The simultaneous presence of injunctive and descriptive SN messages indicating that reporting fake news is a socially desirable behavior and that other people are reporting fake news leads to a higher amount of reported fake news than the use of only one of the two types of SNs.

However, a special characteristic of descriptive SNs is that the strength of a descriptive SN can be varied by adjusting the number of people who show a certain behavior. Current research demonstrates a positive correlation between the strength of a descriptive SN and the influence on behavior (Demarque et al. 2015; Kormos et al. 2014). For example, Dermarque et al. (2015) reveal in an online shopping environment that information about a higher number of people who bought an ecological product leads to more other customers buying ecological products. That is, increasing the strength of a descriptive SN focuses the attention of people on a desirable behavior (represented by the injunctive SN) and motivates them to behave in the desired way (Schultz et al. 2007). Vice versa, a weak descriptive SN indicating that most people do not behave in a desirable way may be counterproductive (Cialdini et al. 1990).

For fake news, we suspect that the presence of a stronger descriptive SN focusing a user's attention on a specific post. At the same time, the descriptive SN is highlighting the injunctive SN and is increasing the likelihood that a user is reporting a specific fake news post. What's even more, we expect that the more users have reported a particular fake news story, the greater is the likelihood that other users will also report this fake news story. However, the results of Wong et al. (2016) show that the intention to report depends on the perceived benefit of the reporting action. Users report more often when they feel that reporting also helps to solve the problem (Wong et al. 2016). Therefore, we suspect that an increase in the strength of the descriptive norm could lead to the user perceiving that there is a little additional benefit in his or her own reporting action since enough others have already reported a specific fake news story to draw the attention of the social media provider. This effect counteracts the positive influence of both the descriptive and injunctive SN. We suspect that from a certain strength onwards, the reduction due to the low perceived additional benefit surpasses the increase due to the descriptive SN. If this threshold is exceeded, a further increase in strength leads to a reduction in the probability of reporting. Based on this, we deduce the following hypothesis:

H 4.3.4 The strength of a descriptive SN – while an injunctive SN is also present – affects the likelihood of a user to report a news story as fake. The influence is non-monotonic and follows an inverted u-shape.

Despite the desired effect of SNs (that fake news is reported), it should be noted that also undesired effects of SNs may occur. Specifically, truthful news might be reported as fake. We do not hypothesize any effects here, as this is not the focus of the present research. However, we include this undesired effect on reporting behavior in our analyses to test whether the hypothesized benefit of SNs comes at a price in terms of misleading reports.

Study 1

To test our theoretical hypotheses, we conducted two online experiments, Study 1 and 2. The studies are closely related to corroborating each other with Study 2 going deeper on H 4.3.2.

Method

Participants: We recruited participants for our experiment via Facebook groups for students. Students are an interesting group as they use social media regularly, in many cases daily (Pempek et al. 2009), which makes them particularly open to the frequent exposure to fake news. Using an online (rather than a lab) experiment and sampling via Facebook are sometimes discussed as methodological weaknesses. For our research context, we rather consider these design choices as strengths supporting the experiment's ecological validity. We recruited participants via German Facebook posts. Hence, the experiment was implemented in German. Screenshots presented in Appendix C.2 are a translation. We motivated participation by a raffle for three vouchers.

To ensure the quality of the responses, we include only participants who use social media at least once a week and who completed the experiment. Explicit checks were made to ensure that participants interacted with all materials of the experiment in order to see all the content. This leads us to data from 320 active social media users. The sample is strongly composed of students (n = 293) with an average age of 23.2 years. Two hundred twenty-four participants are female, 267 participants state that they use social media as a source of information on current events, public affairs, and political issues at least once a week. Previous studies also highlight the importance of social media as a news source (Newman et al. 2017). The participants rarely report posts (mean 0.12 per day, standard deviation of 0.51). The low reporting rate is an indicator in favor of our assumption that individuals often do not report fake news. Appendix C.1 has more details.

Task: We provide the participants with a self-developed news feed so that we can manipulate and monitor the experimental environment. In order to increase the degree of realism and to create a familiar environment for participants, we orient our development towards the best-known news feed design developed by Facebook (Newman et al. 2017). We display content in the form of posts containing a source, an image, a headline, and a lead sentence. We exclude other information that could be used by the participants as potential references to evaluate the truthfulness of a news story according to its general acceptability (social influence bias), such as the number of likes, comments, and shares. In addition, every post has three standard Facebook buttons for interaction with the news feed: a like button, a comment button, and a share button. We have added an additional report button. Such a reporting feature already exists on Facebook, but it requires several clicks to use it. We have simplified the reporting process to reduce possible difficulties and, therefore, to better separate the effects of the different treatments. Figure 10 shows an exemplary implementation of a post in the experiment.

The news feed consists of 15 posts of three different types. First, the news feed includes five fake news posts containing fake news stories that spread widely on Facebook (in German-speaking countries) and are proved to be false (Silverman 2016; Schmehl 2017). We modified the fake news posts based on common fake news characteristics such as spelling mistakes, flashy formatting, or obviously altered images, to make it easier to identify them. Every fake news story is assigned to a non-existent news source to reduce authenticity further. Besides the fake news, the news feed includes five real news posts. These contain no characteristics of fake news and originate from authentic news sources. Besides the news posts, the news feed also contains five neutral posts of imaginary friends to create an environment that is somewhat realistic and, thus, increases the external validity. The display order of the posts is generated randomly for each participant. Appendix C.2 contains all posts.

Treatments: We confront participants with different SNs and observe fake news reporting behavior. We use a 2 (injunctive SN: present vs. absent) x 2 (descriptive SN: present vs. absent) between-subject design:

(1) Control treatment: No SNs were used in the news feed.

(2) Injunctive SN treatment: The injunctive SN was shown at the beginning of the news feed.

(3) Descriptive SN treatment: Five of the ten posts with news content, whether they are fake news or real news, were randomly flagged with a descriptive SN.

(4) Combined treatment: Combination of (2) and (3).

To vary the strength of the descriptive SN, we use a 5-level within-subject design in the treatments with descriptive SNs. Each level occurred exactly once per participant.

Based on random allocation, the number of participants per treatment ranges from 77 to 83. To test the structural equality of the four treatments, we conduct ANOVA tests (for age and usage behavior of social media) and chi-squared tests (for gender, education, and occupation) (Lee et al. 2000). The results indicate no differences at a 5% significance level. Thus, we assume structural equality. A detailed list of the results can be found in Appendix C.3.

Independent variables: To be effective, SNs should be perceived as appropriate and credible (Croson et al. 2009), and they should be present in perception or in the focus of attention (Cialdini et al. 1991). Based on these recommendations, we developed implementations for injunctive and descriptive SNs. We paid particular attention that the norm messages do not appear like an instruction by the experimenter but as part of the social media news feed. The injunctive SN is displayed as a text at the beginning of the news feed (left-hand side of Figure 10). The text reads as follows: "Fake news is increasingly threatening public opinion. It is, therefore, important that our users report such inappropriate content to improve quality for all of us.” The participants have to confirm the text by clicking on the corresponding button in order to see the news feed. Thus, we can ensure that participants recognize the injunctive SN (focus of attention).

We implement the descriptive SN as additional information for a post, indicating how many users have already reported this post (right-hand side of Figure 10). By increasing the number of reports, the strength of the descriptive SN also increases. In our experiment, we distinguish five distinct levels (5, 25, 125, 625, 3125) to cover a broad spectrum without reducing credibility. The assignment of the strength of the descriptive SN to a post is random. Additionally, the number of reports for each post randomly increases by up to a maximum of 10% to ensure that our results are not based on a specific number of reports and that the underlying pattern and prominence of numbers are hidden from the participants.

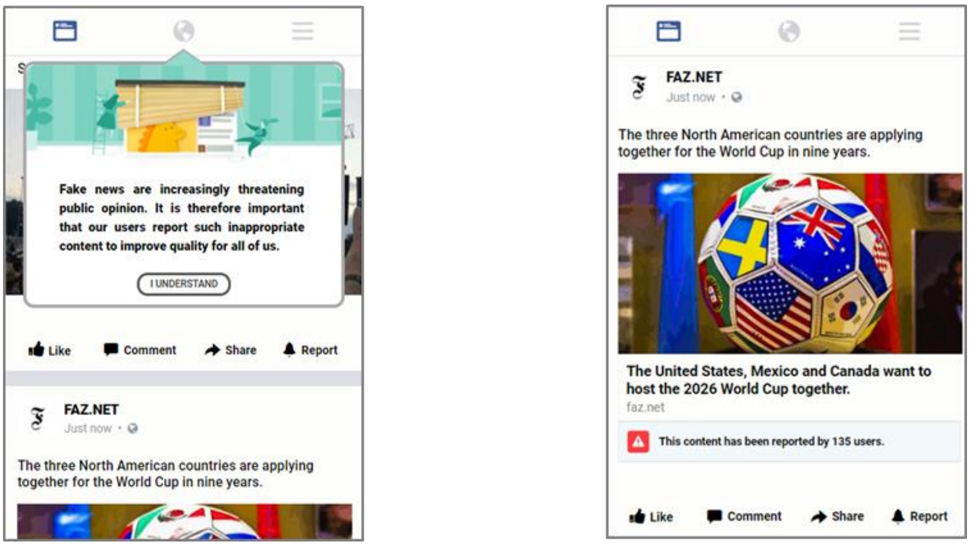


Figure 10: Implementation of the injunctive SN message (left) and a descriptive SN message (right)

Dependent variables: As the primary dependent variable, we measured how much fake news was reported by a person. Additionally, we also measured how much real news was reported by a person. This gives us an indication of how well the fake news was recognizable as such. For statistical tests, we apply a 5% significance level.

Procedures: Participants were informed that the aim was to investigate the behavior of users in social media. We deliberately avoided mentioning fake news. Participants were asked how often they use social media as a news source and how much they trust this information. Subsequently, the interaction phase began, which consists of three stages: the introduction run, the natural interaction run, and the fake assessment run. First, users saw an introductory news feed in addition to an explanation of the four buttons, in order to learn how to use the news feed. Second, in the natural interaction run, the participants saw the news feed with all 15 posts in random order and were asked to interact with it as if they do with their personal news feed. It is only in this phase that the four treatments differ. Next, in the fake assessment run, the participants saw the news feed again, but all interactions were reset. The participants received explicit instruction to report fake news in the news feed and were incentivized by a higher probability of winning a voucher when performing well, specifically in the fake assessment run. Among all this, the natural interaction run is of the highest interest for testing our hypotheses. The fake assessment run should not be overrated for potential effects from the multiple exposure and order of posts as well as the changed incentive as compared to the natural interaction run. We merely use it to assess whether participants were able to differentiate between fake and real news if they paid close attention. As expected, both the fake news and the real news were identified as such by the participants (see Appendix C.4).

The experiment ended with brief questionnaires on social media usage and demographics and with a debriefing session in which the participants were informed about the fact that we had manipulated different features of the posts and that these do not necessarily correspond to the real world. Figure 11 summarizes the process of the experiment. We performed a pre-test (Appendix C.4). Appendix C.5 reports all scales.

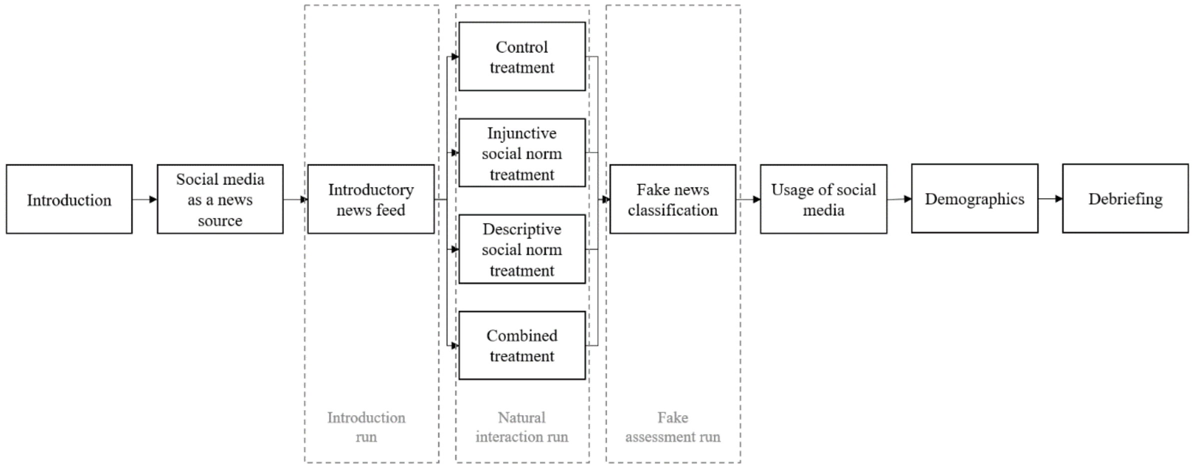


Figure 11: Summary of the procedure of the experiment (Study 1)

Results

In order to measure the effect of SNs on fake news reporting, we compare the number of reported fake news within the four treatments by means of regression analysis. Due to the discrete nature of the dependent variable (integers in the range from zero to five) and non-normally distributed residuals in linear regression, we apply an ordered logistic regression (Agresti and Kateri 2011). The assumptions of an ordinal dependent variable, at least one categorical independent variable, and no multicollinearity are fulfilled by design. In addition, we performed a Brant test and confirmed that the assumption of the proportional odds is satisfied (see Appendix C.6).

To analyze the differences between the control treatment and the SN treatments, we use the control treatment as a baseline. In order to investigate the differences between the combined treatment and the injunctive SN treatment and the descriptive SN treatment, we perform a rebase in which we use the combined treatment as a baseline. Table 11 summarizes the results.

Base Treatment			Coefficient	Standard Error	p-value		Odds Ratio	Related Hypothesis
Control	Treatments	Injunctive SN	0.674	0.326	0.039	*	1.962	H 4.3.1
		Descriptive SN	0.456	0.335	0.173		1.578	H 4.3.2
		Combined	1.373	0.321	< 0.001	***	3.947	H 4.3.3
	Intercepts	0 | 1	0.927	0.249	< 0.001	***		
		1 | 2	1.793	0.263	< 0.001	***		
		2 | 3	2.312	0.276	< 0.001	***		
		3 | 4	3.406	0.328	< 0.001	***		
		4 | 5	4.558	0.451	< 0.001	***		
	Nagelkerke's R^2: 0.068							
Combined	Treatments	Control	-1.373	0.321	< 0.001	***	0.255	H 4.3.3
		Injunctive SN	-0.699	0.292	0.017	*	0.497	
		Descriptive SN	-0.916	0.303	0.003	**	0.400	
	Intercepts	0 | 1	-0.446	0.209	0.033	*		
		1 | 2	0.420	0.209	0.044	*		
		2 | 3	0.940	0.217	< 0.001	***		
		3 | 4	2.032	0.270	< 0.001	***		
		4 | 5	3.185	0.409	< 0.001	***		
	Nagelkerke's R^2: 0.068							

Table 11: Results of the ordered logistic regression to compare the four treatments with respect to fake news reporting in Study 1

Note: In the first model, the control treatment is the baseline, whereas in the second model, the combined treatment is the baseline
+ p < 0.1, * p < .05, ** p < .01, *** p < 0.001; n = 320

H 4.3.1: The effect of injunctive SNs on fake news reporting. The regression analysis shows that the use of an injunctive SN significantly increases the probability of fake news posts to be reported. We observe that, compared to the control treatment, the odds of reporting more fake news increase by 96% (i.e., it almost doubles). Transforming the odds ratio to Cohen‘s d and applying his standard levels for small, medium, and large effects suggest that this effect is significant but small. As we hypothesized, pointing out the importance and the desirability in the social environment of reporting fake news guides the participants’ behavior towards reporting more fake news in their news feed. The results support H 4.3.1. An additional exploratory analysis of the data shows no significant effect on the reporting of real news while applying the injunctive SN (see Appendix C.7).

H 4.3.2: The effect of descriptive SNs on fake news reporting. Comparing the descriptive SN treatment to the control treatment, we see a 58% increase in odds of reporting fake news, but this small effect is not statistically significantly different from zero. H 4.3.2 is not supported.

The data suggest a potential effect of the descriptive SN on the (erroneous) reporting of real news as supposedly fake. The odds increase by 232 % for reporting real news. Yet, this increase also fails to reach statistical significance (p-value of 8%). Details on the regressions for real news reporting are in Appendix C.7.

Not supporting H 4.3.2 is counterintuitive. One reason could be that the descriptive SN was not sufficiently in the focus of attention as – contrary to the injunctive SN – participants did not have to acknowledge it. This idea is supported by the fact that in the ex-post questionnaire, only 25.1% of all participants who saw a descriptive SN state that they used the descriptive SN messages to identify fake news. Another reason could be the different effects of positive and negative descriptive SNs (Cialdini 2003; Cialdini et al. 2006). The news feed contains ten news posts, five real and five fake. In Study 1, the descriptive SN randomly flags five out of these ten news posts as having been reported by other irrespective of whether the news is real or fake. Participants seeing fake news being flagged might perceive this as a positive SN and might report fake news themselves. Participants seeing real news being flagged might perceive this as a negative SN and might either also report real news themselves or might not engage in reporting at all. As the random allocation of the descriptive SN information occurs as a mixed positive and negative SN to most participants, we cannot single out the potentially contradictory effects of positive and negative descriptive SN with the data from Study 1. This is the reason for Study 2 separating the two types of descriptive SNs.

H 4.3.3: The effect of combined injunctive and descriptive SNs on fake news reporting. Comparing the combined treatment with the control treatment, the odds of reporting fake news increase significantly by 295% (i.e., it almost quadruples; large effect). At the same time, the odds of reporting fake news in the combined treatment is also significantly higher than in the injunctive SN treatment (101% increase in odds, derived from the 0.497 odds ration reported in Table 12 for the regression with the combined treatment as a base) and the descriptive SN treatment (150% increase in odds). Thus, the data supports H 4.3.3.

In addition, the data reveals that the number of reports for real news is also significantly higher for the combined SNs as compared to the absence of SNs (control treatment; see Appendix C.7). In the control treatment, three persons reported exactly one real news. In comparison, in the combined treatment, seven participants reported one real news, and five participants reported two real news. For the descriptive SN only, we found an insignificant increase in real news reporting. It appears that adding the injunctive SN amplifies this effect.

H 4.3.4: The effect of the strength of descriptive SNs on the likelihood of reporting news stories as fake. To evaluate the impact of the strength of a descriptive SN on the participants' reporting behavior, we use a logistic regression analysis. The binary dependent variable describes the reporting status of a post indicating whether the post was reported by the corresponding participant or not. The independent variable describes the strength of the descriptive SN. Since we cannot assume a direct connection between the numerical value of the strength of a descriptive SN and the influence on behavior, we interpret the strength of the descriptive SN as an ordinal factor. We focus on the combined treatment since the effects of the descriptive SN are more prominent. Since each of the 82 participants of this treatment has seen five fake news posts, our sample for this analysis consists of 410 fake news. We use the Huber-White procedure to account for clustering in observations that might arise from having multiple posts in a news feed (Huber 1967; White 1980). Table 12 summarizes the results. Appendix C.8 additionally varies the baseline.

The strength of a descriptive SN increases the probability of fake news posts being reported (Table 12). We observe that, compared to no descriptive SN (strength 0), the odds are higher by 106% at a strength of 5, 128% at a strength of 25, 219% at a strength of 125, and 269% at a strength of 625. All these differences to the baseline are statistically significant; the effect sizes are small to medium. We observe indications that the weakest descriptive SN also exerts the least influence. For the strongest descriptive SN (strength 3,125), the odds are lower than for any other strength (see Table 12), statistically significantly lower than for the strength of 625 (Appendix C.8), and not statistically significantly different from the odds for the baseline of no descriptive SN (Table 12). We observe that the probability of reporting a post increases to a strength of 625 and then drops for the strongest descriptive SN tested in our experiment. The ex-ante hypothesized explanation is that with the increasing number of reports from other users, the perceived benefit of one's own reporting action is reduced (Wong et al. 2016), leading to this inverted U-shape. Thus, the results support H 4.3.4.

Strength	Coef-ficient	Standard Error	p-value		Odds Ratio	Related Hypothesis
Intercept	-1.473	0.214	< 0.001	***	-	H 4.3.4
5	0.723	0.313	0.021	*	2.061	
25	0.823	0.389	0.035	*	2.277	
125	1.160	0.329	< 0.001	***	3.190	
625	1.306	0.311	< 0.001	***	3.691	
3125	0.570	0.353	0.106		1.768	
					Nagelkerke's R^2: 0.069	

Table 12: Results of the logistic regression models analyzing the probability of a news post being reported based on the strength of the descriptive SNs

Note: The baseline is a strength of 0
+ p < 0.1, * p < .05, ** p < .01, *** p < 0.001;

Nagelkerke's Pseudo R^2 for this logistic regression is only 0.069. Clearly, other factors beyond the descriptive SN affect reporting. Likely the content, the source, and other post-level characteristics influence reporting along with individual-level differences among participants. As each headline and news source is used only once, our data does not allow for headline or source-effects. In addition, we examined how the strength of a descriptive SN affects the reporting of real news. No significant differences could be observed (see Appendix C.8).

In summary, data from Study 1 supports hypotheses H 4.3.1, H 4.3.3, and H 4.3.4. We do not find empirical support for H 4.3.2 so far. A reason might be that for most participants, the specific implementation of the descriptive SN mixed elements of a positive and a negative descriptive SN. Thus, it is up to Study 2 to further test H 4.3.2.

Study 2

Study 1 tested H 4.3.2 without differentiation between positive and negative descriptive SNs. As mentioned above, seeing fake news posts being flagged by other users can have a different effect compared to seeing real news posts being flagged. Study 2 investigates this further.

Method

Study 2 is an online experiment similar to Study 1, with a notable difference in the treatments.

Participants: Recruiting procedures and filters regarding complete and valid data sets were like in Study 1, only using different Facebook groups to attract different participants (see Appendix C.4 for details). The sample comprises 159 participants who use social media more than once a week. Again, the sample is strongly composed of students (n = 118) with an average age of 25.6 years. Ninety-five participants are female, 143 participants state that they use social media as a news source at least once a week. In total, 146 do not report any fake news or report rather little in comparison to their social environment, for more details, see Appendix C.1.

Task: The task was the same as in Study 1. However, we modified the original fake news to decrease the level of authenticity further. We added emoticons to two of the fake news and changed the term 'German Armed Forces Mission' to 'War mission.' We replaced one fake news completely because it was not well recognized by the participants in Study 1. We replaced the real news stories as they were outdated at the time of Study 2. These measures further increased the difference between fake and real news so that the positive and negative descriptive SN is more readily experienced. See Appendix C.4 for details on the adjustments.

Treatments: We used three treatments: (1) *control treatment*, (2) *positive descriptive SN treatment* (where all five fake news but no real news was flagged with a descriptive SN, with each strength occurring once), (3) *negative descriptive SN treatment* (where all five real news but no fake news was flagged with a descriptive SN, with each strength occurring once).

Independent and dependent variables: Identical to Study 1.

Procedures: The procedures were largely identical to Study 1 with two additions. First, in the fake assessment run, the participants assessed for each news post (on a 5-step Likert scale) how fake or realistic the post appears to them. Second, previous literature on social media users' motivation to report fake news is scarce. To gain further knowledge of it, we asked participants whether they report fake news if they see them. Based on their answer, participants were asked to rate a preselection of possible motivation factors or hurdles on a 5-level semantic differential and to add further factors in a free text. Appendix C.5 lists all survey scales used in Study 2.

Results

Studies 1 and 2 are comparable in terms of reporting behavior: Neither the number of reported fake news nor the number of reported real news is statistically significantly different between the control groups of both studies (Chi-square test, p-values of 0.565 and 0.284 respectively).

In order to analyze H 4.3.2 in more detail, we compare – as in Study 1 – the reporting behavior of fake news of the participants in the different treatments by using an ordered logistic regression. Table 13 summarizes the results. We see that the use of a positive descriptive SN leads to a 37% increase in odds as compared to no SN, but the effect size is less than small, and the effect is not statistically significant. At the same time, the results show that a negative descriptive SN hardly leads to a change in the reporting behavior of fake news as compared to no SN since the odds increase by only 1%, and no significant effect is observed. Overall, even in the extreme setting of Study 2, no influence of the descriptive norm on the reporting behavior of fake news can be detected, which corroborates our results from Study 1.

Base Treatment			Coefficient	Standard Error	p-value		Odds Ratio	Related Hypothesis
Control	Treatments	Positive descriptive SN	0.312	0.367	0.394		1.366	H 4.3.2
		Negative descriptive SN	0.010	0.387	0.980		1.010	
	Intercepts	0 | 1	0.357	0.287	0.196			
		1 | 2	1.052	0.307	< 0.001	***		
		2 | 3	1.668	0.325	< 0.001	***		
		3 | 4	2.007	0.376	< 0.001	***		
		4 | 5	2.628	0.287	< 0.001	***		
							Nagelkerke's R^2: 0.006	

Table 13: Results of the ordered logistic regression to compare the three treatments for fake news reporting in Study 2

Note: The control group is the baseline
+ p < 0.1, * p < .05, ** p < .01, *** p < 0.001; n = 159

In addition, Study 2 provides insights into motivation factors and hurdles of reporting fake news. Seventy-five participants stated they would report fake news if they see them, and 84 would not. Summarized, the motivation factors and hurdles to reporting are mainly intrinsically motivated and altruistic and not because of possible financial incentives. The answers from the free text have a large overlap with the preselected factors. However, two further hurdles could be identified. Table 14 shows the ranked results.

Motivation Factors and Hurdles	Mean	Median	Standard Deviation
Preselected motivation factors to report fake news			
I would like to avoid negative consequences that result from the dissemination of fake news	4.667	5	0.553
It is important to me to help other people to form their opinion based on true facts	4.427	5	0.808
A correct news landscape is important to me	4.347	5	0.846
I would like to improve the living conditions for myself and other people	3.880	4	0.986
I do not want that the relevance of the topic will be reduced by the dissemination of fake news	3.800	4	1.078
Fake news causes an unpleasant feeling in me	3.747	4	1.079
I would like to improve the quality of the social media platform I use	3.507	4	1.234
I want to share my knowledge with others	2.960	3	1.083
I hope for appreciation from my social environment	1.813	2	0.982
I expect a material or financial incentive	1.707	1	1.037
Preselected hurdles to reporting fake news			
I do not believe that the act of reporting has any effect or counteracts the dissemination of fake news	3.774	4	1.144
I am only a consumer of news and do not actively participate in the public discourse	3.702	4	1.138
The procedure of reporting is too complex for me	3.131	3	1.377
Reporting fake news offers me no material or financial incentive	2.762	3	1.444

I am not interested in public opinion	2.690	3	1.270
Fake news serves for entertainment	2.333	2	1.205
I've never seen fake news	2.155	2	1.227
Fake news is subject to the freedom of opinion	2.107	2	1.006
I don't know how to report fake news	2.012	1	1.266
Further hurdles to report fake news (answers from free text question)			
Lack of responsibility		-	
Own uncertainty		-	

Table 14: Motivation factors (n = 75) and hurdles (n = 84) to report fake news measured on a 5-level scale

Note: 1 corresponds to "strongly disagree" and 5 to "strongly agree"

Discussion

Fake news is a greater threat than ever before, especially with the widespread and fast dissemination on social media. We aim to contribute to the emerging body of IS research on fake news by studying socio-technical interventions as remedies for fake news (Bernard et al. 2019). Specifically, we theoretically and experimentally analyze the effect of injunctive and descriptive SNs on social media users' reporting of fake news.

Recent IS research investigated individuals' ability to detect fake news and the underlying factors affecting cognition and judgment (Moravec et al. 2019). A key result was that confirmation bias plays a strong role, and many individuals are unable to distinguish real from fake news (Moravec et al. 2019). A complementary piece of recent IS research explored the effect of source ratings (e.g., user source ratings derived from aggregating users' article ratings) as a remedy against fake news on social media (Kim et al. 2019). Here, a key result was that information on the credibility of a news source displayed alongside a social media post affects individuals' belief in the news (Kim et al. 2019). Our research is complementary to these two recent papers: Both concern individual users' ability to detect fake news. Detection is a prerequisite for subsequent reporting. Inversely, the reporting we study is a means to obtain user source ratings, which then, in turn, improve detection. These different mechanisms and remedies may reinforce each other.

Our work leverages knowledge on fake news in social media and knowledge originating from social psychology, specific theories on SNs, and the focus theory of normative conduct. Our results suggest that SNs do have an impact on the reporting behavior of users in online environments. Social media users can be guided in their behavior by highlighting desirable behavior and making transparent what other users are doing. However, caution is needed, as SNs can also give rise to the undesirable behavior of reporting more real news as *by-catch*. However, as the reporting levels of real news are substantially lower than the reporting levels of fake news, we argue that the overall effect of improving reporting by a combination of different SNs messages is beneficial.

Researchers working on fake news on social media have to consider the effects of SNs in the future. Researchers engaging in behavioral research should integrate SNs in their theorizing and should examine the prevalent SNs in their samples to better understand their behavior. Researchers engaging in design science research should be aware of the power of SNs to steer user behavior. They may leverage this power in their socio-technical designs and should be cautious of the side effects of invoking SNs unintentionally. Furthermore, to rigorously evaluate their designs, researchers should keep in mind that SNs can be an explanatory variable of user behavior in social media environments and, thus, might measure SNs to control for their effect. Finally, it is important to remember that publishing about phenomena like reporting behavior in social media also signals a descriptive SN. Our motivation to investigate the effects of SNs in social media is based on the fact that only a few people report fake news today, which itself is a descriptive SN. In doing so, we also explain that most users in social media do not report fake news, and therefore we actually highlight undesirable behavior. Scholars should consider this effect when publishing their results. But from a positive point of view, our work also highlights the injunctive SN that reporting fake news 'ought to be done' (Cialdini et al. 1990).

In more detail: First, we find an indication that SNs are not always an effective approach to guide social media users' reporting decisions towards reporting fake news. On the one hand, injunctive SNs are a motivational tool highlighting moral evaluation in social media environments, which is comparable with similar studies in this field (Rimal and Real 2003; LaBrie et al. 2010; Neighbors et al. 2007; Wenzel 2004; Park and Smith 2007; Baumgartner et al. 2011). On the other hand, empirically, we do not find a significant effect when we apply isolated descriptive SNs in social media environments. This result is contrary to the findings of other researchers that descriptive SNs are an appropriate tool to motivate desirable behavior by describing the behavior of others (Cialdini et al. 1991). The combined application of both – injunctive and descriptive – SNs has already led to the strongest effects in previous studies as it affects different motivations (Cialdini 2007; Bernedo et al. 2014; Cialdini 2003; Schultz et al. 2008). We demonstrate that this is also the case in the domain of reporting fake news in social media. Following the focus theory of normative conduct (Cialdini et al. 1990), descriptive SNs seem to focus users' attention on the injunctive SN (when this is sufficiently present) and thus increase the probability of users' reporting fake news posts.

Another argument for the combined application of injunctive and descriptive SNs is more practical. Our injunctive norm message has been displayed when the news feed was loaded and then clicked away by the user. This procedure is not very user-friendly, so a social media provider would not display the message every time the platform is accessed. However, the descriptive SN constantly reminds the user of the injunctive SN, thus redirecting the user's focus back to what is 'ought to be done' (Cialdini et al. 1990).

Second, we provide further evidence that descriptive SNs can also backfire. Users in social media environments decrease in reporting fake news when the strength of a descriptive SN raises a threshold. One explanation may be that users lose their motivation to report fake news due to the reduced benefit, as a significant number of users have already reported the message. As Wong et al. (2016) already showed, the intention to report depends on the perceived benefit of the reporting action. This does not fit with other studies that show a positive correlation of behavior and strength of a descriptive SN (Demarque et al. 2015; Kormos et al. 2014). These studies, however, examined effects in scenarios such as online shopping (Demarque et al. 2015), which promises a different individual benefit than reporting fake news in social media.

In addition, we also looked at whether SNs have a negative effect on the reporting of real news. Although our data do not show a significant effect, we cannot exclude that with increasing strength of a misplaced descriptive SN, real news becomes more likely to be further reported as fake news. For this reason and the fact that strong descriptive norms reduce the reporting of fake news, one may react from a certain threshold and choose alternative approaches. For example, starting from a defined number of reports, the exact number could be hidden, and an article or source rating could be shown.

The results have implications for practitioners who design social media platforms or interfaces. First, to encourage users to report fake news, messages should be displayed, highlighting the injunctive SN that this behavior is socially desired. Second, in order to further increase the number of reports for fake news, practitioners should consider a combination of injunctive and descriptive SNs to focus the attention of their users. In this, the use of descriptive SNs can also have negative effects. The result is a trade-off between a higher number of reports for fake news and the accuracy of reporting fake or truthful news, in which one has to decide according to a social media platform’s specific objectives, design, users, and environmental influences.

Nevertheless, our findings are subject to limitations. The generalizability of our empirical results is limited by our design of the experimental materials (inspired by Facebook, one wording for the injunctive SN, one way of displaying the descriptive SN) and by our participants (young, German, educated, active social media users). Furthermore, in practice, users of social media platforms cannot always be asked to confirm the same injunctive SN. A variation of the wording may help to confront the user more often with injunctive SNs and to promote desirable behavior. The motivational factors and hurdles for reporting fake news that we identified in Study 2 may support the development of injunctive SN messages. The composition of the posts’ headlines, the perceived reputation of the posts’ sources, the posts’ images and subjects, as well as the users’ familiarity with the news, might all affect the identification of fake news. We did not explore these factors related to detection but solely focused on reporting behavior.

There are several implications for future research. First, future research might empirically test our hypotheses with more diverse or complimentary samples in other forms of social media that use other design elements or have a different structure, such as forums or instant messengers. In this context, future work should also deal with the boundary conditions of the descriptive SN in more detail. Our results suggest that there is an inflection point beyond which descriptive SNs no longer improve user behavior. Such inflection points should be examined in further empirical studies. Going beyond artificial experimental settings to field experiments on live social media platforms would greatly strengthen external validity. Second, future work should explore whether SNs from the direct social environment have stronger influence (as prior research from other domains suggests (Baumgartner et al. 2011; Neighbors et al. 2007)). Also, researchers should examine how SNs and factors affecting cognition and judgment interact (e.g., familiarity with news sources, topics, or headlines (Kim et al. 2019; Moravec et al. 2019)). Third, it appears a promising direction to study the effect gamification on fake news reporting, that is “[...] the use of game design elements in non-game contexts” (Deterding et al. 2011, p. 10) or “a process of enhancing a service with affordances for gameful experiences in order to support user’s overall value creation” (Huotari and Hamari 2012, p. 19). Gamifying fake news reporting could take, for example, the form of leaderboards or point systems. Fourth, SN-based interventions are appropriate socio-technical tools to foster active user engagement, which allows for follow-up measures such as user source rating or third-party fact-checking. Hence, we suspect that descriptive SNs also improve the ability to detect fake news. Future research may investigate this effect. Lastly, it is not only the users‘ responsibility to take action against fake news. Providers of social media platforms are also obliged (by the public or new laws (Oltermann 2018)) to search for additional mechanisms to reduce fake news.

Overall, this chapter investigates the influence of SNs on users reporting behavior in the context of fake news. We could theoretically derive and empirically validate that SNs can cause behavioral changes by guiding users towards reporting fake news. The potential for the application of these mechanisms seems considerable, as existing social media applications do not have to be adapted at great expense but can simply be extended by adding SNs. Therefore, they are a promising tool to support existing methods against fake news and thus reduce their dissemination and negative impact.

5. Design Knowledge from a Technocentric Perspective

IT enables structural, communicative, and decision-making changes in organizations (Sarker et al. 2019; Pinsonneault and Kraemer 2002). Concerning sustainable development, business enterprises, as a dominant form of social organizations, contribute to the damage and burden of our environmental nature (Melville 2010). IS have been one of the main reasons for increasing productivity in recent decades (Gholami et al. 2016). However, IS also have the potential to enable sustainable transformation in companies and to achieve environmentally sustainable businesses (Gholami et al. 2016; Melville 2010). For instance, appropriate information allows assessing the status quo and deciding for suitable solutions that foster sustainable improvements (Gholami et al. 2016). IS scholars criticize that green IS research primarily develops theories for description. Instead, IS research should develop solution-oriented design theories that contribute to achieving a more sustainable lifestyle (Vom Brocke et al. 2013). A promising starting point to foster sustainability is more efficient value creation processes to avoid waste (e.g., Martínez-Jurado and Moyano-Fuentes 2014). Waste, as a result of inefficiency, occurs not only with regard to unnecessary overuse of resources, overstocking, or generation of garbage due to unused resources. Waste also occurs with unexploited resources of time or inaccurate or unavailable information leading to false decisions. Therefore, IS artifacts can make use of the increasing penetration of digital technologies and the associated digitalization to avoid such waste. Chapter 5.1 in the following outlines the need for advanced methods and presents a new approach to analyze and design lean, digitally supported value creation processes in companies and organization. Subsequently, Chapter 5.2 evaluates the utility, quality, and efficacy of the newly introduced approach by means of two exemplary applications to avoid waste in hospital logistics processes. Major parts of Chapter 5 conform with Heger et al. (2020a), Heger et al. (2020b), Denner et al. (forthcoming), Heger et al. (forthcoming), and Thim and Heger (forthcoming).

5.1. Value Stream Modeling and Notation (VSMN) – Developing a Domain-Specific Modeling Language

Digital transformation poses great opportunities and challenges for science and business (Gimpel et al. 2018; Legner et al. 2017; Vial 2019). With the increasing availability of digital technologies and the aspired vision of fully networked and virtualized entities in organizations, many new developments and potentials arise. Further information flows created by digital networking can provide the required information for tactical and strategic decision making in business processes and organizational routines at the right time and in the right place. Besides, networking makes it possible to control, for example, production and logistics processes more efficiently and flexibly (Emmrich et al. 2015; Kelkar et al.).

A current example: The COVID-19 pandemic puts society at large, the healthcare sector, and especially hospitals, including their intensive care units, and hospital staff under specific pressure. In the wake of the pandemic, the need for more tightly integrated material and information flow becomes even more obvious than before. Among others, increased short-term demand for individual articles (e.g., face masks) posed major challenges for the supply of materials. There is still a lack of transparency about existing stocks, which made efficient redistribution to the point of demand difficult. In bed logistics, as another example, there was a lack of transparency regarding available beds and ventilators, which made it nearly impossible to react quickly. As soon as bottlenecks occur, such as in Italy or Spain, this can lead to life-threatening waste of time.

Examples from practice and academic literature help to illustrate the potential of digitalization more generally: As part of a research project within two hospitals (A and B), we identified different types of waste in their logistic processes even before the outbreak of COVID-19. In addition to common types of waste, such as unnecessary movements of material due to inappropriate processes, we often identified various kinds of waste that can be avoided by appropriate information flows and improved information availability. An example is the lack of transparency about stock levels in decentralized warehouses. Existing information is often not used, which can lead employees to work and partly fill buffers without prioritizing orders (known elsewhere) and not meeting existing demands directly. Furthermore, various media breaks slow down processes: Some information is printed out on paper to be manually re-entered into the system elsewhere. These examples are only a selection of the waste that the targeted use of digital technologies and the design of appropriate information flows can avoid.

While all these examples relate specifically to logistics processes within the hospitals involved, evidence from interviews with experts suggests that similar problems can be found in production and logistics processes in other domains and industries like the automotive or commerce sectors. The experts confirmed that the increasing availability of digital technologies and the associated information networking could solve many of the mentioned problems and raise potentials in today's value creation processes (see Appendix D.3 for details). The opportunities for reducing waste and achieving lean production and logistics processes by applying digital technologies are also supported by scientific literature. Uckelmann (2014), for example, summarizes eight opportunities for value creation processes that arise from the application of digital technologies and increasing information flows. Among others, these are the avoidance of latency, the elimination of manual data acquisition, and higher information availability. Furthermore, Metternich et al. (2017) outline the reduction of uncertainty through up-to-date information, providing information at the point of value creation, as well as the possibility of dynamic inventory adjustments.

So how can processes be designed in such a way that waste is avoided and continuous value creation – that is, a value stream – is created (Klevers 2009)? The results of a competing artifact analysis in both disciplines, lean management and IS research, show that there is no appropriate DSML (Frank 2013). Notably, there is no DSML, which – in the context of the analysis and design of value streams – allows the investigation of information flows in consideration of durations and lead time. Yet, such a DSML would be beneficial to better use digitalization for capturing the potential for value creation sketched above. For this reason, this chapter aims to

develop a domain-specific modeling language for analysis and design of value streams under consideration of temporal dependencies in information flows and information availability.

Therefore, we apply the IS design science research paradigm (Hevner et al. 2004; Peffers et al. 2007). Our design research results in the VSMN. We specifically foresee users from logistics and production planning at the interface to IT. Analyses (actual state) and designs (target state) of the value stream facilitate the communication between model builders and model users. A representation of the actual and target state permits the collection of recommendations for action, the derivation of a transformation roadmap, and input for the development of suitable IT systems.

The chapter's structure is as follows: first, we derive design objectives and review extant artifacts competing with our to-be-designed VSMN. Next, we describe the research process. Subsequently, we present a metamodel and the graphical notation of VSMN. Lastly, we evaluate VSMN and conclude with a discussion.

Design Objectives and Existing Approaches

The term *digitalization* characterizes diverse socio-technical phenomena and processes of adopting and using digital technologies in broader individual, organizational and societal contexts (Legner et al. 2017). Caused by and contributing to digitalization, companies experience a controlled transformation. This digital transformation includes the introduction of digital technologies to change business concepts, improve existing working processes, create new value streams, and ensure value creation (Gimpel et al. 2018; Vial 2019). For instance, digital technologies are the popular SMACIT technologies (social, mobile, analytical, cloud, IoT) as well as platforms (Vial 2019; Sebastian et al. 2017), or new technological trends such as artificial intelligence, blockchain, or 3D/4D printing (Gimpel et al. 2018). A value stream is the combination of all activities needed to bring a product from the source material into the form required by the customer. Value streams include not only the production processes in the material flow but also the activities with which processes of the material flow are controlled, including the information flow (Klevers 2009).

Design Objectives

For the digital transformation of value streams and the raising of the associated potentials, suitable methods and modeling languages are required. Domain-specific approaches provide multiple advantages. First, a DSML consists of concepts that represent domain-level knowledge. Therefore, modelers and model users do not need to develop or transfer such concepts. Second, a DSML contains domain-specific constraints that must not be enhanced manually and, thus, contribute to model integrity (Frank 2013). It is vital to take into account potential users of the specific DSML to develop a suitable DSML (Frank 2013).

Overall, we aim at designing a DSML for the analysis and design of value streams under consideration of temporal dependencies in information flows and information availability. Potential users are practitioners from lean management, such as production managers, logistics managers, or warehouse managers. However, with the increasing application of digital technologies, practitioners from IT departments will be more involved in the future. This holds true for scholars in research projects regarding lean, digitally supported value streams. For this reason, a suitable DSML must, first and foremost, allow efficient communication between interdisciplinary users through conceptual models. Visualizing such models supports communications among modelers and model users (Wand and Weber 2002). Besides, conceptual models help to analyze domains, design processes, and to document requirements (Wand and Weber 2002). The digital transformation of value streams requires a suitable DSML to consider temporal dependencies of information flows and information available. More precisely, processes are likely to deviate when process participants (e.g., employees) do not have access to relevant information (König et al. 2019). Delayed information availability or, even worse, missing information can lead to delays and thus significantly increase the lead time (Roser et al. 2014; Roh et al. 2019; Kuhlang et al. 2011). The lead time describes the duration of a distinct good to pass the whole process and is one of the most important key figures in designing efficient value creation processes in production and logistics (Bertagnolli 2018). Information flows and information availability do have a direct effect on the lead time and, thus, value streams. In this context, it is particularly important to consider the time dependency when the required information is available in a process.

Thus, in the context of the digital transformation of value streams, a suitable DSML for value stream analysis and design enables the visualization of information availability and temporal dependencies, to recognize the wasted time and to arrange appropriate information flows.

DO 5-1: DSML must enable the visualization of temporal dependencies of information flow and information availability regarding value streams.

Inappropriate information flows and lack of information availability lead to obvious and unnecessary waste in existing processes and, thus, reduce value creation. We also observed such waste throughout a research and development project within two hospitals. We gained insights into current waste due to missing or inappropriate information flows. Experts from other industrial sectors also confirmed these insights. For instance, too little information on actual stocks leads to overstocking, which is problematic for a land shortage, tied-up capital, and potentially perishable goods. Besides, stock shortages (production stoppages) occur due to missing information on actual demands. Without appropriate information, this may also lead to wrong decisions being taken due to uncertainty. That is, corresponding process steps are not carried out at all or with the wrong priority, which leads to unnecessary work.

Furthermore, the value creation processes are often controlled by paper documents and supplemented selectively by IT systems, which slows down the processes and, hence the lead times. These conditions also lead to media breaks and input errors. Ultimately, not all the necessary information is available. For example, data on the expiration date is only recorded sporadically and is not systematically used during the further course of the process. This missing information leads to the use or disposal of expired products.

Against this background, VSMN must enable the analysis of waste in value creation processes due to missing or delayed information. Such waste includes isolated data handling and missing communication, frequent media breaks, and changes between digital and analog media. Those lead to additional effort in the affected process steps and may have effects on subsequent (waiting) processes (Erlach 2010).

DO 5-2: DSML must enable value stream analysis for time wasted due to missing information flows or missing information availability.

The introduction of digital technologies in the course of digital transformation can improve existing working processes and create new value streams (Gimpel et al. 2018; Vial 2019). The digitalization of production and logistics processes, in particular, enables efficiency improvements of internal processes (Metternich et al. 2017). The increasing spread of digital technologies allows new information flows, increases the availability of information, and helps avoid waste. The technologies enable more efficient data acquisition, communication, data processing, and process control. Thus, waste such as unnecessary movements, corrections, or communication breaks can be prevented (Uckelmann 2014). Further examples are the reduction of uncertainty in planning and support systems through up-to-date information (in real-time), providing information at the point of value creation, as well as the possibility of dynamic inventory adjustments, and new possibilities for problem-solving with the help of process data (Metternich et al. 2017).

On the other hand, no new types of waste must arise from information processes. These may be superfluous information movements and collection, but also unnecessary activities (e.g., through manual entry), waiting times (e.g., during data processing), or a flood of data (Uckelmann 2014). The provision of accurate information is, therefore, an important aspect and requires the design of value streams with suitable information flows and information availability under consideration of the dimension time.

DO 5-3: DSML must enable value stream designs taking into account appropriate information flows and information availability.

Competing Artefacts in Lean Management

Lean management has various philosophies and methods to analyze, design, and thus improve logistics and production processes. The aim of lean management is, among others, to maximize value creation by reducing waste of time (Bertagnolli 2018). Comparing different methodological approaches (Appendix D.1) regarding the factor time, the Association of German Engineers (VDI) guideline 2870 (VDI 2012) provides eight methods and their corresponding domain-specific modeling approaches. These guidelines of the VDI are recognized as practice-oriented regulations for maintaining quality standards in many technical areas, including lean management (VDI 2020). We analyzed the eight methods regarding the design objectives. These methods are process standardization, shop floor management, idea management, one-piece flow, quick changeover, value stream method (VSM), U-layout, and Kanban. As a result, the VSM appears to be the most suitable method among the ones suggested by the VDI. In the VSM, value streams are divided into the material flow – describing the stream of physical goods through logistic and production processes – and corresponding information flows, which describe the provision of information. The aim of using this method is a holistic and transparent visualization of processes (Koch 2015). The economic optimization and efficient design of processes lead to an orientation towards customer needs and the reduction of waste in the production process (Bertagnolli 2018). Generally, there are seven classic types of waste: overproduction, overprocessing, unnecessary transports, inventory, unnecessary motions between process steps, waiting times, and bad quality goods (Klevers 2009). Besides, other types of waste, such as unused employee knowledge, may occur in value streams (Bertagnolli 2018). Another type of waste due to the digital transformation of processes and the increasing associated networking of information due to the application of digital technologies affects the processes of order processing and information flows (Erlach 2010). VSM is a practical approach for process optimization to avoid waste and foster value creation and processing times. Reduced processing times result in various potentials like increasing flexibility on changing customer requirements, and reduction of capital commitment in the decrease of stocks (Bertagnolli 2018).

However, a disadvantage of the VSM (especially concerning our DO 5-1, DO 5-2, and DO 5-3) is the lack of consideration and transparent design of time dependencies in the information flow. The notation of the VSM represents application systems, and associated information flows without dependencies on material flow and the dimension time. Thus, the consideration of information flows is one main disadvantage of VSM regarding the increasing availability of digital technologies. As a result, the VSM cannot reflect the advantages and possibilities of digital transformation (Hartmann et al. 2018).

Extensions of VSM were developed to overcome these disadvantages. Meudt et al. (2017) introduced a method for analyzing value streams and information flows (VSA 4.0), systematically mapping the opportunities to derive improvement. Hartmann et al. (2018) built on this and presented the value stream method 4.0 (VSM 4.0), which includes VSA 4.0 and supplements the method through an approach to design digitally supported value streams (VSD 4.0). Therefore, they linked the information flow between a process, storage systems, and data usage. While Hartmann et al. (2018) present a method to analyze and design a value stream and propose a visualization approach, they did not rigorously develop a DSML to enable the visualization of value streams, and its temporal dependencies to information flow. First, the authors did not apply a research method to develop a DSML, as proposed by Frank (2013). In doing so, the authors, for instance, do not develop a comprehensive DSML (cf. Wand and Weber 2002). Second, the contribution of Hartmann et al. (2018) does not achieve our DO 5-1 and DO 5-2, and in part also our DO 5-3: The visualization and analysis of information flow concerning time dependencies, availability, and waste regarding value streams.

The same applies to the extended value stream mapping method, which was introduced by Busert and Fay (2019). The authors developed a six-stage procedure to apply control methods. They harmonized information flows, which is the alignment of information quality to achieve the requirements and conditions of a process step. Busert and Fay (2019) extended the notation of the VSM to allow the visualization of information flows in more detail. Three parallel lanes for a process, information, and information processing enable the modeling of the quality of information required in the process steps. However, Busert and Fay (2019) also do not apply a specific research method for DSML, do not present an abstract syntax, nor do they achieve our DO 5-1, DO 5-2, and DO 5-3.

In summary, VSM is a widely known starting point. Still, neither the original VSM nor its recent extensions enable the analysis and design of dependencies in the value stream to respective information flows and information availabilities. Thus, they fail to deliver upon our DO 5-1 and DO 5-2 to visualize and analyze time dependencies and waste in value streams due to inappropriate information flows or information availability. Merely, the extensions, in part, allow designing value streams and appropriate information flows. However, the extensions do not fully achieve our DO 5-3.

Competing Artefacts in Information Systems

IS research provides different generic modeling approaches for the modeling of time-dependent relationships. Formal approaches, such as Petri nets (Petri 1962; van der Aalst 2013), make it possible to model processes and time correlations, but they are frequently perceived as unsuitable for recording processes for analysis and design due to their complexity (Meyer et al. 2005). Visual approaches, such as the Unified Modeling Language (UML), seem more suitable for many scenarios. For example, UML contains interaction diagrams such as time diagrams to represent state changes or different states of a structural element over time but does not allow visualizing value streams (Object Management Group 2015).

Business Process Model and Notation 2.0 (BPMN) is the de-facto standard modeling language in process management (Allweyer 2015). The specification contains three different types of diagrams: process and collaboration diagrams, choreography diagrams, and conversation diagrams (Allweyer 2015). BPMN permits the visualization of processes using a flowchart, in which activities are linked to decision points (gateways) and sequence flows (Object Management Group 2015). Further, BPMN makes it possible to map the information flows between processes (messages), the data used, and the source systems (data stores) (Object Management Group 2015). Another advantage of BPMN is the opportunity to visualize data-related dependencies (Allweyer 2015). In this way, it can be modeled that different activities depend on the available information.

But BPMN is restricted when it comes to visualizing processes and information dependencies in terms of duration and temporal availability. BPMN does enable the visualization of time dependencies (time events) to illustrate which information is required to execute subsequent process steps appropriately. In contrast, BPMN does not take into account temporal information availability nor visually associates the information flow to the lead time. For this purpose, commercially available software tools for process simulation have been developed, such as Adonis[7]

[7] https://de.boc-group.com/adonis, accessed on 01.11.2019

or Signavio[8]. These tools allow the analysis and design of business processes taking into account time dependencies, but do not represent a visual approach.

To sum up, until today, no rigorously developed DSML exists, which enables the analysis and design of value streams in the sense of material flows and its dependencies to information flows under consideration of the dimension time (and, thus, meeting DO 5-1, DO 5-2, and DO 5-3). Developing a suitable DSML is most promising in our eyes by extending the existing notation of the VSM, which for our domain, is the modeling approach with less need for adaptation as compared to BPMN.

Research Method

A DSML consists of three main components: abstract syntax, concrete syntax, and semantics (Wand and Weber 2002; Nordstrom et al. 1999). The abstract syntax represents a language syntax that defines the structure of the modeling language in the form of a metamodel and maps concepts and rules of the DSML (Gonzalez-Perez and Henderson-Sellers 2008). The concrete syntax provides graphical notation elements required for visualization. The semantics describe how the constructs and relationships of the abstract syntax can be interpreted (Bouhdadi et al. 2007).

The development of the VSMN is based on the design science research methodology (Peffers et al. 2007). Design science research is particularly suited to build and evaluate design artifacts like constructs, models, methods, and instantiations (March and Smith 1995; Hevner et al. 2004). Figure 12 summarizes our research process.

In a first step, we illustrate the relevance of the problem. We derive a need for an advanced DSML from the insights of a research and development project and statements from interviews with experts from different organizations. Besides, we verify the necessity with the help of insights from literature (see previous subchapters).

Next, we derived design objectives (DO 5-1, DO 5-2, and DO 5-3) from literature and insights from the research project. These design objectives are domain-specific requirements. Subsequently, we ran two focus groups with nine scientists to discuss the meaningfulness of the design objectives. The first focus group was attended by six business IT specialists with expertise

[8] https://www.signavio.com/, accessed on 01.11.2019

in process management and BPMN. Three logistics and lean management specialists with expertise in VSM and design participated in the second focus group. In both groups, we described the initial situation and outlined the purpose of VSMN. In summary, both groups stressed the need for a suitable DSML against the background of its purpose. Our subsequent competing artifact analysis supports this.

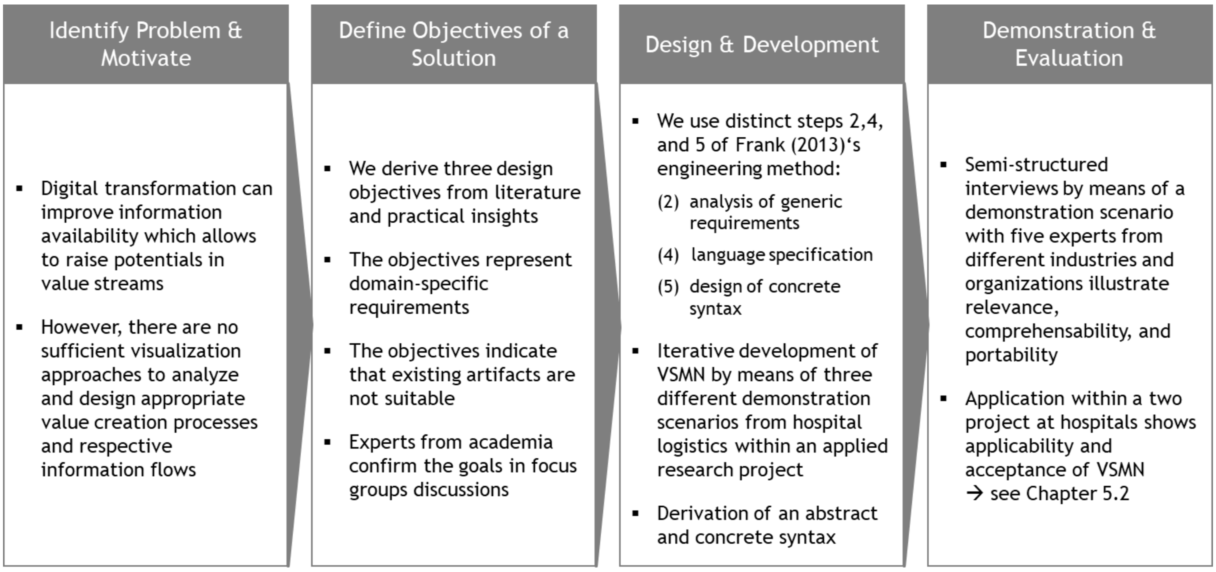


Figure 12: Research process based on Peffers et al. (2007) with selected steps from Frank (2013)

The abstract syntax (and subsequent concrete syntax) originated from an iterative search process (Hevner et al. 2004). Since Peffers et al. (2007) do not specify the design and development of DSMLs in particular, we follow Frank's engineering method for DSML (Frank 2013). VSMN was developed iteratively, analogous to steps 2, 4, and 5 in Frank (2013): analysis of generic requirements, language specification, and design of graphical notation. The overarching design science research process already covers the other steps of Frank (2013) (clarification of scope and purpose, analysis of specific requirements, and evaluation and refinement). During the development, we considered the design objectives and the generic requirements for DSML, according to Frank (2013): familiarity, invariability, level of detail and abstraction, and mapping of language concepts. We also consider further design science research evaluation criteria for design artifacts, which are practical relevance, understandability, ease of use, and applicability (Sonnenberg and Vom Brocke 2012) to demonstrate utility, quality, and efficacy of VSMN (Hevner et al. 2004).

To develop VSMN, we modeled several logistic processes of the hospitals (A and B) being part of the research and development project. We started with the common process of *goods receipt*, which is not hospital-specific. We then refined the constructs and relationships and extended the modeling to other, more complex processes (bed logistics, material supply in hospitals wards). For clarity and brevity, Appendix D.2 summarizes the iteration steps.

Subsequently, we evaluated VSMN using semi-structured interviews with five practitioners from manufacturing companies, logistic consulting, and a healthcare organization to ensure the requirements and design science research evaluation criteria. In doing so, we also ensured the portability of VSMN to other organizations than hospitals. Throughout the interviews, we introduced the abstract and concrete syntax of VSMN and discussed it using the *goods receipt* scenario. The experts' feedback further enhanced VSMN so that we can present the mature version of VSMN in this chapter. Appendix D.3 provides the *goods receipt* scenario and a summary of the expert's feedback and comments. Throughout Chapter 5.2, the thesis presents two applications of VSMN in the scope of projects with hospitals.

Specification of the Modeling Language

In the following, we present the abstract syntax of VSMN as a UML class diagram. Then, we introduce the concrete syntax, which consists of the graphical notation elements.

Abstract Syntax

We present a semi-formal metamodel (Frank 2013) using UML as the de-facto standard in object-oriented modeling (Lodderstedt et al. 2002). The metamodel has a modular structure to facilitate the understanding of the abstract syntax. After the presentation of the core model, the process objects and flow objects are detailed as partial concepts. For this purpose, simplified UML class diagrams are used to specify the metamodel (Eriksson et al. 2013).

The *core model* consists of four central constructs: value stream, process step, flow, and digital technology (Figure 13). The value stream represents a high-level construct and consists of a sequence of process steps (e.g., varnishing) and flows (e.g., transport of goods). It begins with a source – this being an internal (e.g., warehouse) or external (e.g., supplier) starting point of the value stream – and ends with a sink – this being an internal (e.g., subsequent value stream) or external (e.g., customer) endpoint of the value stream (cf. Klevers 2009). Furthermore, a value stream consists of at least one process step, as well as at least two flows. A flow always connects exactly two process steps, including the source and sink. A process step is followed by at least one flow. Process steps and flows have a duration (e.g., processing time, transport time). The lead time of the higher-level value stream summarizes the duration of the individual process steps and flows and measures the time required by a good in the material flow from the source to the sink.

Besides, digital technologies (Vial 2019; Gimpel et al. 2018) can be used both to execute process steps and flows. These technologies enable the acquisition of data, communication, and networking of entities, data processing, and control via corresponding actuators in value streams (Uckelmann 2014). Digital technologies have an interface that allows the transmission of information to and from the technology to a process step or flow.

A *process step* is an abstract superclass that contains all process elements (Figure 14) and corresponds to the *FlowNode* class in BPMN (Object Management Group 2011). Each process step can be executed by one or more participants (people who execute activities are not necessarily employees, cf. Alter 2013) or digital technologies (e.g., automated guided vehicles). Each process step is assigned to exactly one location (e.g., warehouse, work station, etc.). There are three types of process steps: processes and activities, gateways, and events.

The central process steps are the processes and activities, whereby processes generally consist of different activities (Object Management Group 2015). The individual process steps can be regarded as activities analogous to the BPMN notation. Depending on the granularity, distinct activities can be considered in turn as separate processes, including their activities, allowing different degrees of abstraction and detail. We include processes and activities simultaneously as activities can also constitute steps, which are poorly defined and do not have an inner sequence (cf. Alter 2013). In the following, we use the term *activity* synonymously like *process*.

Activities are subdivided into material-oriented and information-oriented. While material-oriented activities depict physical tasks and contribute to the material flow, information-oriented activities depict all tasks to create or process information.

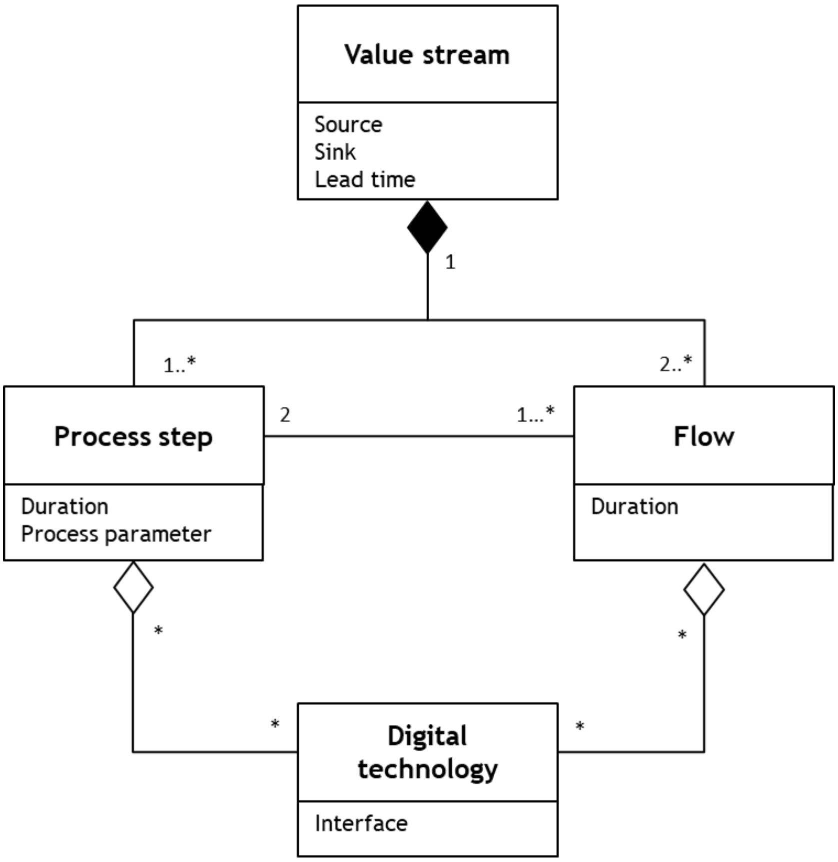


Figure 13: Core model representing the abstract syntax of VSMN

In addition to activities, gateways are a form of a process step. Similar to BPMN (Object Management Group 2011), gateways are used to represent branches. On the one hand, information-oriented activities can be executed in parallel. On the other hand, information can lead to different decisions in the process flow (logical and exclusive or).

Events (Object Management Group 2015) are process steps that occur during the value stream. For instance, events exist when a dependency occurs between the information flow and the material flow (see *flow metamodel*, *transfer*). Here, information is required to execute a material flow.

A *flow* links two process steps and fulfills two tasks (Figure 15). Similar to *SequenceFlow* elements in BPMN (Object Management Group 2011), flows represent the logical sequence of the process steps and, thus, the process flow. Besides, flows represent either the material flow or the information flow and have been adopted from VSM (Klevers 2009).

Material flows are divided into push and pull in the same way as in the VSM notation. If the previous activity moves the material to the next activity (follow-up activity), push arrows are used. If the subsequent activity requests the material as required, the material is pulled and visualized in the value stream with pull arrows. Also, resources, such as forklifts, pallet trucks, or outer packaging, can be used to facilitate the flow.

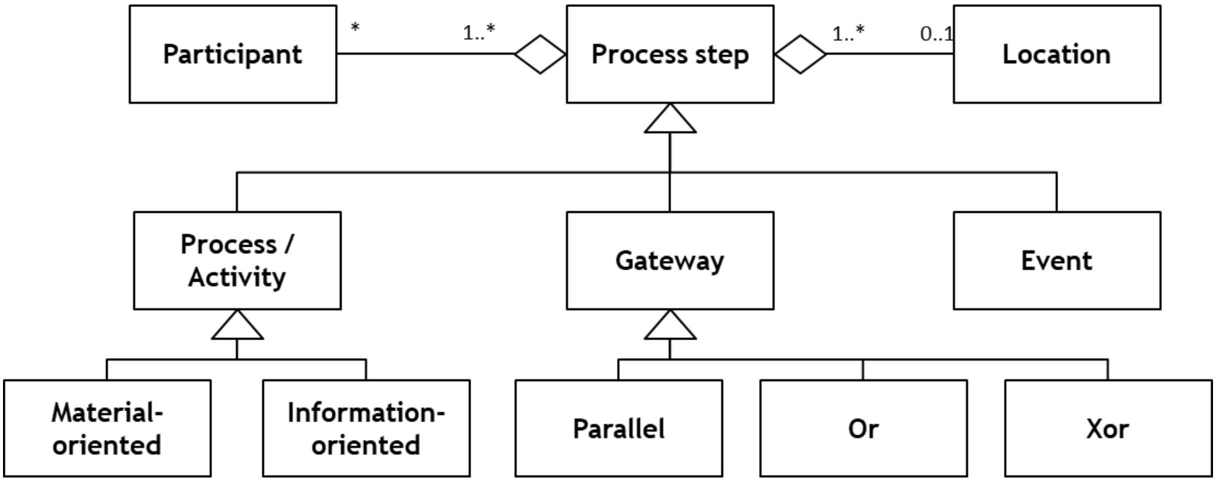


Figure 14: Detailed representation of the abstract syntax of the process steps

An information flow arises in the logical sequence of information-oriented activities. The information flow consists of data, which can be analog or digital. A special case of flow is a transfer. Transfers always take place when an activity in the material or information flow triggers a subsequent activity in the other flow. These transfers must be given explicit attention both in the value stream analysis and in the value stream design, which is why a separate notation element is introduced for this purpose. Insufficient transfers can lead to waste, for example, by leading to waiting states or wrong decisions under uncertainty.

Buffers can arise in any (material or information) flow that holds a stock of material or information and lead to waiting times. In the information flow, buffers represent missing or existing but unused information that can lead to delays and waste. In the material flow, buffers represent the physical accumulations of material. The duration in which, for instance, an article lies in the buffer, is called the resting phase.

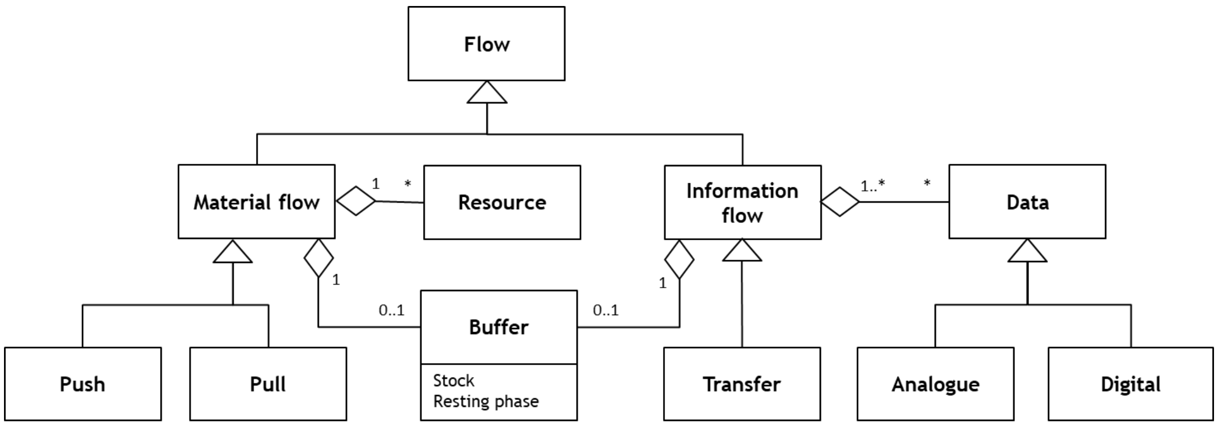
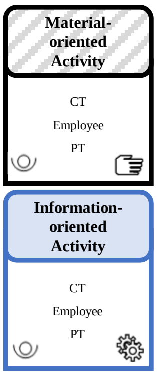


Figure 15: Detailed representation of the abstract syntax of flows

Concrete Syntax

In the following, we present the concrete syntax of our VSMN being the graphical notation elements. For most elements, we can (re-)use elements from the VSM notation following Klevers (2009) and Womack and Jones (2013), and from the BPMN notation following the Object Management Group (2011). This approach promises intuitive usability and increased familiarity for modelers and model users. Besides, we created new elements for the few concepts that do not exist within either notation. To proof familiarity, we highlighted these elements throughout our evaluation steps.

The VSMN notation includes elements for process steps, flows, and digital technology. Table 15 lists the elements for process steps. The activity symbols contain the name of the respective process step, which is shown within the box header. Underneath the header, the process parameters of the activity are listed in the body. Typical process parameters are the cycle time, the number of operators, and the processing time. This selection is only a part of the possible process parameters in VSM notation. For example, parameters like the change over time, pack sizes, and product variations can also be used for production processes (Womack and Jones 2013). Besides, task type elements, which we take from BPMN, describe whether the activity is executed manually or automatically. The material-oriented activity and information-oriented activity are differentiated using differentiation areas to create a visual distinction. Each activity is assigned to a location within the organization, which is highlighted by a grey box. An event visualizes a process signal, which initiates the following activity. Furthermore, source and sink are visualized using a factory symbol following the VSM notation.

Element	*Symbol*	*Description*	*Source(s)*
Process/ Activity		• Material-oriented activity (upper symbol) • Information-oriented activity (lower symbol) • Name of the process step in the header • Enumeration of process parameters (see process parameter) in the body • Task type element (see task type) in the body	• Adapted from VSM (Klevers 2009)
Process parameter (selection)	CT  PT	• Cycle time (CT) (1st symbol) • Number of employees (2nd symbol) • Processing time (3rd symbol)	• Adapted from VSM (Klevers 2009)
Task type	 	• Automatically executed activity, e.g., by digital technologies or machines (upper symbol) • Manually executed activity, e.g., by employees (lower symbol)	• Adapted from BPMN (Object Management Group 2011)

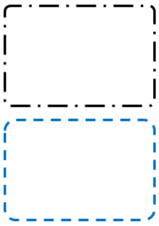
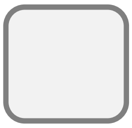
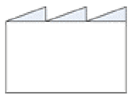
Differentiation area		• Differentiation of the material flow (upper symbol) • Differentiation of the information flow (lower symbol)	• Self-developed
Location		• Location within the company, e.g., goods receipt	• Following BPMN swim lanes (Object Management Group 2011)
Event		• An event in the material flow or information flow to start a process	• Adapted from BPMN (Object Management Group 2011)
Source/sink		• Suppliers • Costumers of the product/service	• Adapted from VSMN (Klevers 2009)

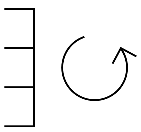
Table 15: Elements of process step notation

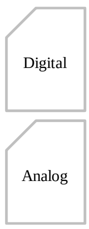
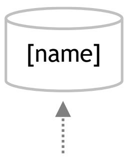
The activities are visualized with flow elements connecting them to display entire process sequences. First, different material flows between process steps can occur (Table 16). Between material-oriented activities, the striped arrow symbolizes the pushing of material to the following activity. In contrast, the symbol for a pulled material flow is a withdrawal in combination with an icon of a supermarket (Klevers 2009). The latter is a symbol for a warehouse from which material is picked when required. Resources are visualized using symbols such as a truck. All notation elements for material flows so far are taken from VSM notation.

Second, between activities, two different arrows are used to visualize the transmission of information. On the one hand, between two information-oriented activities, there is a continuous arrow according to a sequence flow of BPMN. On the other hand, a dashed arrow with a dot at the beginning visualizes the transfer of information between material-oriented and information-oriented activities. This arrow is newly created and has been validated throughout our evaluation steps. Each arrow should contain the transferred data.

Third, necessary information within the value stream could be stored in or retrieved from digital technology, which is, for example, a platform (e.g., enterprise resource planning system) or merely an IoT-device. A dotted arrow is visualizing the interface, according to message flows in BPMN.

Finally, buffers are visualized between the different process steps through a triangle symbol. In some cases, the material flow may not go on without the right information at the right time. The material rests in the buffer without any transfer or movement when information is missing or not available yet. To visualize this temporal expansion and to connect the material flow visually, we introduce a resting phase symbol, which has been validated throughout our evaluation.

Element	*Symbol*	*Description*	*Source(s)*
Material flow (push)		• The material flow between two process steps • Transfer of material via push movement	• Adapted from VSM (Klevers 2009)
Material flow (pull)		• Warehouse (supermarket) with defined stock of products, the subsequent activity picks up the material as required (left symbol) • The material flow between two process steps (right symbol) • Transfer of material via pull movement	• Adapted from VSM (Klevers 2009)

Information flow		• Information flow between two activities (upper symbol) • Transfer of information between the material and information flow (lower symbol)	• Adapted from BPMN (Object Management Group 2011) or self-developed (lower symbol)
Data		• A data object, digital, e.g., order data (upper symbol) • A data object, analog, e.g., delivery note on paper (lower symbol) • A short description of the data	• Adapted from BPMN (Object Management Group 2011)
Digital technology		• Digital Technology, which acts as a data source or data store (upper symbol) • A digital technology can be a classic SMACIT technology (social, mobile, analytical, cloud, IoT) and new technological trends such as artificial intelligence or wearables (Vial 2019; Sebastian et al. 2017; Gimpel et al. 2018) • Interface to/from a digital technology (lower symbol)	• Adapted from BPMN (Object Management Group 2011)
Buffer		• The stock of material or information between two process steps	• Adapted from VSM (Klevers 2009)

Resting phase		• Visual supplement for the representation of the temporal expansion, while the material or the information remains in a buffer and waits for process steps in the respective other flow (information or material)	• Self-developed
Resource	 (example)	• Resources are tools or aids that are used in the material flow • There is no exhaustive list of symbols; here a truck is listed as an example	• Adapted from VSM (Klevers 2009)

Table 16: Notation of the flow elements

Below the value stream, there are different timelines to visualize the lead time (Figure *16*). The timeline is separated into two different areas. The upper timeline shows the durations within the information-oriented activities and flows. The lower timeline is assigned to the durations within the material flow (as known from VSM). Both durations are summed separately at the end. The summations allow a separate evaluation of the information-oriented and material-oriented flow, which is new compared to VSM.

The recording of the duration of separate activities, events, and flows allows a more detailed analysis of the lead time. Table 17 summarizes the different durations.

Abbreviation	*Description*
t_{PTmat}	Processing time of material-oriented activities
t_{PTinf}	Processing time of information-oriented activities
t_{PTpar}	Processing time of an activity (material- and information-oriented), which is executed simultaneously to another activity (auxiliary variable)
t_{Tmat}	Duration of material flow (i.e., transportation), including resting time in buffers
t_{Tinf}	Duration of information flow (i.e., transmission), including resting time in buffers
t_{Tpar}	Duration of a flow, including resting time in buffers (material- and information-oriented), which takes place simultaneously to another flow (auxiliary variable)
t_E	Duration of an event in material or information flow (belongs to waiting time)
t_{ET}	Duration of information transfer between the material and information flow (in both directions, also belongs to waiting time)

Table 17: Summary of duration variables to calculate the lead time

Generally, the processing time is the duration it takes for an individual good (or information) to be processed within an activity (Klevers 2009). A sink in the time axis visualizes both types of processing times (material-oriented and information-oriented). Furthermore, there are waiting times that summarize the duration of a material or information flow, including resting time in a buffer (Klevers 2009). At the end of the value stream, all the durations are added up, categorized into waiting times in the upper part of the box and processing times below. Subsequently, both sums of the duration of the information flow and material flow are shown separately. The separation has the advantage that both durations of the different areas can be evaluated independently. The last part is the sum of all the times, again divided into waiting and processing times. In the case of parallel activities or flows, simultaneous durations are summed up into auxiliary variables and are deducted from the lead time.

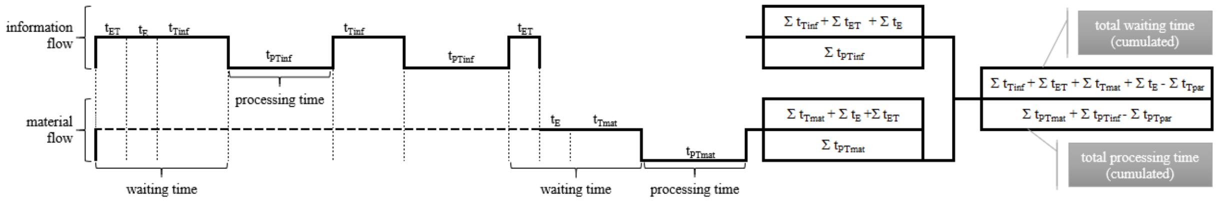


Figure *16*: Timeline of a value stream based on Klevers (2009)

The total of all durations – waiting, processing, material, and information – results in the lead time. VSM and also VSMN aim to improve the value stream coefficient, which is the ratio of total processing time to lead time. The optimal state persists when the processing time is equal to the lead time. In that case, no waste or waiting times occur within the value stream, and the value stream coefficient reaches zero (Bertagnolli 2018; Tomanek and Schröder 2018). The overall aim is to analyze and designing integrated information and material flows. As such, the lead time and integrated value stream coefficient are important. The additional calculation of the value stream coefficient for information flow and material flow separately allows a better understanding of potential inefficiencies.

The total of all durations – waiting, processing, material, and information – results in the lead time. VSM and also VSMN aim to improve the value stream coefficient, which is the ratio of total processing time to lead time. The optimal state persists when the processing time is equal to the lead time. In that case, no waste or waiting times occur within the value stream, and the value stream coefficient reaches zero (Tomanek and Schröder 2018; Bertagnolli 2018). The overall aim is to analyze and designing integrated information and material flows. As such, the lead time and integrated value stream coefficient are important. The additional calculation of the value stream coefficient for information flow and material flow separately allows a better understanding of potential inefficiencies.

Evaluation

To evaluate the generic and domain-specific requirements, as well as design science research evaluation criteria, we obtained feedback on VSMN from experts working in different industry sectors.

When we obtained a suitable version of VSMN, we presented it to five experts from different sectors. Therefore, we carried out semi-structured interviews with five practitioners who are all experienced in both logistics and digitalization. The interviews aimed to evaluate VSMN in terms of the requirements, evaluation criteria, and also to ensure portability of VSMN to use cases outside of hospitals. The sectors included manufacturing companies and service providers (Table 18).

Organization	*Company type*	*Employees*	*Revenue [Euro]*	*Job title*
Production I	manufacturer of technical springs	~160	~100 m	Production manager
Production II	Manufacturer of gearboxes	~2,300	~500 m	Senior Digitalization Manager
Production III	Automotive manufacturer	~130,000	~100 bn	Digitalization Manager
Service I	Logistics consulting	1	information not available	Managing Director
Service II	Hospital (S)	~2,000	information not available	Logistics Manager

Table 18: Anonymous overview of the experts in the interviews

In the interviews, which lasted about 60 minutes, we first motivated the topic and presented our design objectives. We then introduced VSMN using the abstract and concrete syntax and demonstrated it using the *goods receipt* scenario as an example. Subsequently, we discussed with the experts whether VSMN complies with our design objectives and the generic requirements. We also investigated where VSMN could be applied in the companies. Appendix D.3 provides the *goods receipt* scenario and an overview of the experts' feedback and comments.

First, the experts confirmed that the practical relevance was given. Second, the interviewees agreed that our VSMN was understandable, had a high degree of familiarity due to its construction on known notations, and possessed a sufficient level of detail (latter both relevant for ease of use). In particular, the expert from *production I* – he is a production manager – discussed the meaningfulness of the notation element for the resting phase and the separate visualization of material and information flow. During this, we discovered that the resting phase was not included in the metamodel and added it accordingly. Furthermore, the expert identified different processes within his company, which were likely to display waste as a result of insufficient information flows and availability. In particular, he stressed that VSMN made it possible to visualize media breaks in the sense of unnecessary switching between analog and digital information.

The second expert is a senior digitalization manager of *production II*. He spoke more about the applicability and understandability from the point of view of the IT department. He pointed out how, with the help of VSMN, it was possible to reveal which unnecessary dependencies exist between material and information flow. The transfer arrows make the dependency visible and could - from his point of view - be reduced to a minimum during the design of the value stream. In his opinion, transfers were sources of waste as they were likely to cause delays (waiting for information or information that was available but unused) or to be prone to errors (e.g., when entering data).

The third expert (*production III*) primarily discussed the metamodel and the concrete syntax. As a result, we have changed the names of the construct *Environment* to *Location* and *Means of Transport* to *Resource*. *Environment* refers to the physical environment in which a process step takes place. However, confusion with the software environment is possible, so we changed the name. *Means of transport*, on the other hand, had been too narrowly defined and was renamed to take account of the outer packaging, for example. We have also revised the activity symbol from a colored to a patterned distinction. If models using VSMN are printed out, the copies are usually in black and white, so that it is not possible to distinguish between the colors.

The coloring of the symbols was also a subject of the discussion with the expert from *service I*. Being a logistics consultant, he criticized the coloring since red (which we used for information flows at that point) is usually used to identify a problem. Furthermore, he stressed the need for a clear definition of a source and a sink. That is why we have made our explanations in Section 4.1 more precise. Regarding our metamodel, the expert noted that an additional differentiation of the buffers into material-oriented and information-oriented was not necessary, as the distinction did not provide further information. We removed this from the metamodel accordingly to reduce complexity. Finally, we discussed the visualization of the timeline and durations with the expert. Before the interview, the total times were not divided into those of the material and the information flow, to be able to evaluate the totals individually. Besides, it is now possible to differentiate the duration when parallel process steps or flows occur and to determine which element has a longer duration and thus contributes to waste.

Finally, the fifth expert is a logistics manager at yet another hospital (S) (*service II*). He emphasized that the VSMN is very well suited to compare process variants with each other and to recognize the differences. Also, the illustration makes it possible to see which process steps, especially in the information flow, can be moved in advance or combined in order to reduce the processing time. As a disadvantage, the expert complains that, in his opinion, VSMN is too complex to communicate the processes to employees in general. In the discussion, the expert and the authors emphasized that the VSMN should primarily support logistics managers and production managers in the analysis and design of value streams and communication with IT managers. We agreed that other modeling languages such as BPMN could be used for communication with a broad audience.

Nevertheless, based on the discussion, we have simplified the representation in some symbols to achieve less complexity. Before the interview, there were several symbols for data being transferred (for instance, between material and information-oriented activities or data that is received from a digital technology). Additionally, there was also a distinction between digital and analog data. Finally, we also limited VSMN to one symbol for events. Previously there was a distinction between starting, intermediate, and end events, which was not necessary.

In summary, it can be stated that – according to the expert interviews – VSMN fulfills the generic requirements (familiarity, invariability, level of detail and abstraction, and mapping of language concepts) and the design objectives (being domain-specific requirements). Moreover, the discussion with experts from different industry sectors also shows that VSMN is achieving the design science research evaluation criteria (practical relevance, understandability, ease of use, and applicability) but also its portability to other industries.

Discussion

Digital transformation enables new value creation processes and the improvement of existing workflows (Gimpel et al. 2018; Vial 2019). The application of digital technologies allows us to design existing value streams more efficiently and to eliminate waste (Uckelmann 2014). This chapter presents the *Value Stream Modeling and Notation*, which is based on elements from VSM and BPMN. The VSMN allows an analysis of material flows and information flows under consideration of temporal dependencies and enables the analysis and design of value streams. In addition to a concrete notation, we present a metamodel that represents the abstract syntax of VSMN. Subsequently, we evaluate VSMN based on semi-structured interviews with experts from five different industry sectors.

From a theoretical perspective, VSMN contributes to existing design knowledge about DSMLs and expands the field of VSM by a new approach for mapping material flows, information availability, and information processing while considering temporal dependencies. Being a modeling language, VSMN is a valid design artifact in accordance with Hevner et al. (2004). DSMLs, in general, are models as they are statements about constructs and their respective relationships (March and Smith 1995). In doing so, VSMN is a level 2 contribution and an improvement following Gregor and Hevner (2013), as VSMN is a new solution for a known problem that cannot be solved using existing approaches.

Different modeling languages and extensions of the VSM already exist. Modeling approaches from the field of business process management (e.g., BPMN, Petri Nets) and IS research (e.g., UML) are not suitable approaches for analyzing and designing material flows, and information flows in value streams, considering the time dependencies. Other lean management methods (e.g., total productive maintenance) besides the VSM do not focus the factor time sufficiently, too. Extensions of the VSM merely concentrate on the method itself and lack the development of a concrete DSML. Busert and Fay (2019) provided an extension of the VSM to take into account the possibilities of digital transformation. Their method includes a modeling approach to visualize information quality aspects. Meudt et al. (2017) and Hartmann et al. (2018) also introduced a further developed VSM 4.0, which included a new approach for value stream analysis and value stream design. Besides, they proposed a modeling approach to link information flows between process steps and a storage system. However, they did not suggest a comprehensive modeling language as, for instance, suggested by Wand and Weber (2002) or Frank (2013). VSMN, in contrast, is a complete DSML, which is including an abstract syntax, concrete syntax, and semantics. It can fill the gap of a missing modeling language for analysis and design of information flows and information available in value streams under consideration of temporal dependencies. Thus, VSMN achieves our DO 5-1, DO 5-2, and DO 5-3. Table 19 compares VSMN and related approaches regarding the design objectives.

	DO 5-1	*DO 5-2*	*DO 5-3*
BPMN (Object Management Group 2011)	(✓)	-	(✓)
VSM (Klevers 2009)	-	(✓)	(✓)
Busert and Fay (2019)	(✓)	(✓)	(✓)
Meudt et al. (2017)/Hartmann et al. (2018)	(✓)	(✓)	✓
VSMN	✓	✓	✓

Table 19: Comparison of VSMN and related approaches

Note: - = not fulfilled; (✓) = partially fulfilled (mostly missing time dependency); ✓ = fulfilled

Our research is beset with some limitations: Due to the iterative approach taken, we cannot formally claim that all possible scenarios and constructs of the VSMN have been covered. However, because of the positive evaluation by five experts from different companies, we are convinced that VSMN covers the most important constructs and scenarios. Nevertheless, future work may validate and extend VSMN. With the evaluation, we show that VSMN addresses an important problem and represents a meaningful solution. We also show that VSMN meets the design science research evaluation criteria. However, our evaluation does not yet include an application of VSMN to analyze and design logistic processes. The thesis overcomes this shortcoming in Chapter 5.2. From a methodological perspective, a modeling language is presented, not a modeling method. In particular, the suitability of the previous VSM (Klevers 2009) and its extensions (Hartmann et al. 2018; Busert and Fay 2019) should be questioned against the background of VSMN and further developed if necessary. At last, the development of VSMN is based on the notation of VSM. However, also other process optimization methods or modeling languages, such as BPMN, for instance, could be promising starting points leading to ultimate design solutions, too.

Nonetheless, VSMN presents a practical approach for visualizing value streams that consist of material flow and respective temporal dependencies to information flows and information availability. On the one hand, VSMN enables researchers and practitioners to examine existing value streams to analyze their process steps and (material and information) flows for waste based on insufficient information flows or information availability. The dependencies between information flow and material flow can be displayed with the help of transfers and events. Also, the transmitted data can be visualized to detect media breaks. VSMN represents a practical approach to support the design of value streams. As a result, VSMN allows depicting value streams in a way that suitable information flows and availabilities become visible to reduce lead times. VSMN facilitates communication between model builders and model users and supports the development of appropriate IS and the efficient use of digital technologies in production and logistics processes.

5.2. Designing Lean, Digitally Supported Logistics Processes for Hospitals – Case Studies to Evaluate VSMN

Up to this point, VSMN has only been evaluated using expert interviews. The application took place during the development of VSMN, but there was no application in practice. This chapter presents two cases to carry out a further evaluation. First, this chapter presents results from the research and development project in which the VSMN also has been developed. Among others, the research project aimed at developing a reference model for lean, digitally supported material logistics processes in hospitals. The results originate from a process analysis in two hospitals (A and B). The reference model constitutes a value stream design using VSMN. Subsequently, the proposed model has been evaluated throughout a prototypical implementation. Second, this chapter presents the results of another applied-research project with the hospital (M). In this project, the material logistics processes of the cardiac catheter laboratories were examined for existing waste using VSMN. Also, we developed a target state for the respective material logistics processes. Together, both applications in practice demonstrate the applicability of VSMN. At the same time, the results also constitute prescriptive knowledge for material logistics in hospitals from a technocentric perspective.

Material Logistics in Hospitals

Logistics is a basic building block in hospitals. The core business of hospitals is to ensure the medical care – diagnostics, treatment, and nursing – of patients. However, due to the special character of medical services, the success of treatments is based on a targeted interaction between the primary value-adding actors (e.g., patient, doctor, nursing staff) and the supporting service and logistics processes (Kriegel et al. 2016). Hospital logistics comprises the planning, storage, and provision of physical goods that are necessary for the treatment of patients (e.g., beds, drugs, or sterile treatment instruments) (Volland et al. 2017). Hospital logistics aims to ensure that the right person (patient, doctor, or nurse) is in the right place at the right time and that the necessary materials are available in the right quality and quantity at the correct cost. Hospital logistics is a portfolio of transport, storage, picking, and information services. There are numerous logistics processes, which differ in particular concerning contact with patients, the degree of individualization, and product groups. Due to high-quality requirements, patient individuality, and the need for short-term reactions, particular challenges arise for hospital logistics.

These challenges include the availability and transparency of information about materials and equipment involved in the process, as well as continuous process flow across wards and departments. Despite a considerable variety of suppliers and products, the reduction of waste (e.g., time, material, employee resources) and the establishment of an understanding of production and value creation are essential to achieve cost savings by avoiding process inefficiencies or overcapacities in stock.

As an important part of hospital logistics, material logistics ensures the management of the hospital with medical consumables and care materials. Today many of the challenges mentioned above have not been addressed sufficiently, as the following two cases demonstrate. Typical waste are, for example, unnecessary material movements due to inappropriate processes, high material stocks due to uncertainty, often unused information, and media breaks, which slow down the processes. As already stated throughout Chapter 5.1, such waste can be avoided by the targeted use of suitable digital technologies. Digital technologies allow achieving transparency to prevent, for instance, overstocking and, thus, obtain cost-efficient material logistics. In the following, the doctoral thesis at hand presents two examples from research projects that demonstrate the avoidance of waste using digital technologies.

Developing a Reference Model for Lean, Digitally Supported Material Logistics Using VSMN

The research and development project *Hospital 4.0*[9] aims to improve hospital logistics regarding quality and efficiency. The question at hand is how to design logistics processes to address the growing challenges posed by changing needs (e.g., growing patient numbers) and rising opportunities from the increasing penetration of digital technologies. To this end, the research and development project was concerned with two logistic processes, one of which is material logistics.

The project aims to provide the answers obtained in generic form. Therefore, we derived a reference model that describes a target picture of material logistics using VSMN. Reference models have their origin in data management (Legner et al. 2020). The underlying process of reference modeling includes all actions that are necessary to create and apply reusable models

[9] This research and development project is funded by the German Federal Ministry of Education and Research (BMBF) within the “Innovations for Tomorrow’s Production, Services, and Work” Program and implemented by the Project Management Agency Karlsruhe (PTKA).

(in terms of reference models) (Fettke and Loos 2004). Reference models are abstract representations of domain-specific knowledge, which include descriptive and prescriptive knowledge for socio-technical problems (Schermann et al. 2009; Legner et al. 2020). Reference models play a key role in addressing comprehensive problems related to increasing digitalization (Legner et al. 2020). Reference models can capture and describe existing problems or transport scientific statements (Fettke and Loos 2004). In the context of this chapter, reference models are understood as a set of general statements that are valid for a class of companies or, in particular, hospitals (Fettke and Loos 2004). A reference model consists of a generic design of a reference process (or several sub-processes), which represents a blueprint for a process in hospitals.

The development of a reference model always follows a similar procedure (Fettke and Vom Brocke 2019): Definition of the application domain, construction of the model elements, evaluation of the reference model, adaptation, and verification (Fettke and Loos 2007; Fettke and Vom Brocke 2019). Within the scope of the research and development project, the existing material logistics processes in the participating hospitals (A and B) form the starting point for generalizing and abstracting the reference models. Therefore, we first present selected insights of the process analysis. Subsequently, we introduce the reference model, which includes two reference processes for *goods receipts* in the central warehouse of a hospital and *goods handling in decentralized storage locations* (e.g., ward storage). We applied VSMN to design both reference processes. Finally, we provide insights into the prototypical implementation of a software instantiation.

Analysis of Existing Material Logistics Processes

In the beginning, we collected qualitative and quantitative data about the material logistics processes. Therefore, we interviewed employees of both hospitals (A and B), who handle the material logistics, and carried out observations of the actual processes. Also, we collected data from the materials management (MaMa) system to make quantitative statements. In the following, we provide a brief overview of the two processes in focus at the hospital (A).

Goods receipt: The process with the parcel service or the assigned forwarding agent bringing packages or pallets to the goods receipt area. The logistics employee confirms the receipt of the goods on the handheld of the parcel service's employee. At the hospital (A), approximately 100 deliveries per day are accepted in a two-hour time window.

Subsequently, a logistics employee processes the delivered items one by one. The employee processes the goods in order of his[10] own prioritization. Therefore, the delivery note is removed, which is attached to the outside of the parcel in more than 80% of cases.

In case of a missing delivery note, the logistic employee makes a complaint to the supplier directly by telephone or e-mail. The employee retrieves the respective contact person from the MaMa system. A subsequent delivery note will be accepted in the aftermath.

The hospital's purchase order number is available as a barcode on approximately 2% of the delivery notes. Thus, most of the time, the employee has to enter the data manually into the MaMa system. Subsequently, he also enters the delivery note number to link it to the purchase order entry. He compares the items and quantities on the delivery note manually with the purchase order. The actual quantity of the items is not checked consistently. On average, the employee is working two minutes to process a package from picking up until it is ready for storage or shipping.

Special cases represent under- or overdelivering. In the case of under-delivery, the delivered items are posted into the MaMa system. The physical difference remains in the system as an outstanding delivery. In the case of over-delivery, the employee contacts the purchasing department by telephone. Together, the departments determine whether to accept the items or to make a complaint.

After the delivery note has been completely processed, the employee notes the current date, signs the note, and places the items on a designated tray. During a morning tour, an employee of the purchasing department takes the delivery notes from the central warehouse to the purchasing department, where the notes are archived physically.

Stock items are stored in the central warehouse. Therefore, the employee prints out a warehousing receipt from the MaMa system, which is attached to the items. Pass-through items are delivered to a decentralized storage location directly. For this purpose, the employee prints out a hospital-internal delivery note from the MaMa system. The internal delivery note contains the destination, which is also written on the packages. Subsequently, the employee places the items on a transport trolley for hospital-internal shipping.

[10] In order to improve readability, this chapter generally uses masculine form only for personal designations which refer to women, men and other persons, e.g. "his" instead of "his or her". However, this is not intended to express gender discrimination or a violation of the principle of equality.

Pass-through items, as well as stock items, must have been requested by a supply assistant responsible for the material supply of a decentralized storage location. The items are delivered to decentral storage locations either manually or automatically using an automated transport system.

Goods handling in decentralized storage locations: the second process in focus starts upon arrival of the items**.** A light signal indicates that a new delivery has arrived. The supply assistant stores the items into respective compartments. Subsequently, the supply assistant turns the cards of the compartment to the white side (indicating that the corresponding demand request is fulfilled, see below). However, the supply assistant does not check the items' expiration date. This check is carried out twice a month manually.

The supply assistant is responsible for the replenishment of items in the patient rooms as well as the trolleys, which are used as mobile storage during medical care services. Therefore, the supply assistant picks items from the decentralized storage locations. The nursing staff also picks items from the storage location when required. The replenishment of the decentralized storage locations themselves follows a KANBAN procedure (pull principle, see Klevers 2009). The procedure requests required items when reaching a defined point of reordering. The point of reordering is defined according to the historical consumption of each item by the supply assistant manually. The nursing staff or the supply assistant notes an actual demand by turning the cards of the respective compartment to the red side.

The supply assistant records the demand at fixed points in time per week (mostly twice a week). He uses a hand scanner, which must be synchronized with the MaMa system in advance. The supply assistant captures the item identification number from each (red) card. Capturing the identification twice or more often signalizes increased demand. For example, the supply assistant captures the item's identification number three times to demand three internal units. Subsequently, he turns the card of each compartment upside down to signal a captured demand request. At last, the supply assistant transmits the data to the MaMa system in order to generate a demand request. In some cases, the supply assistants also enter demand requests manually to avoid effort for the synchronization. The demand requests are mostly printed out and archived to allow inspection on delivery.

Overall, both processes in focus reveal areas for improvement. The goods receipt process, for instance, requires manual data input and visual comparisons. Also, the process does not ensure the verification of actually delivered goods. The goods handling process, for another example, does not allow stock transparency and requires a complex turning of cards. The reference model, therefore, provides target states for both reference processes aiming to address these areas for improvement.

Reference Model for Material Logistics

The reference model originates from a value stream design using VSMN and includes two reference processes for both goods receipt and goods handling, as introduced above. An important base for the reference model to achieve transparency is the newly introduced unique device identifier (UDI) (European Parliament 2017). The UDI is a machine-readable and human-readable code that makes a medical item identifiable and also contains item-specific information such as a batch number or expiration date. Both reference processes make use of the UDI.

Goods receipt: The first reference process comprises the delivery, entry, and posting of items. The reference process ends before storage (in the case of stock items) or shipping (in the case of pass-through items). The reference process aims to ensure an efficient goods receipt and simultaneous allows quality assurance. Besides, item-specific data (expiration date, batch number, serial number, etc.) is collected at the time of receipt, which can be used at various points in the later course of the material logistics process. The reference process is preceded by a purchasing order from the hospital's purchasing department and a corresponding provision of the items by the supplier.

The reference process starts with the delivery by a supplier or an authorized forwarder. The supplier or forwarder transports the ordered items to the hospital. The supplier previously notifies and optionally registered the delivery in a booking system for time window control of goods receipt. When the items arrive, the logistic employee of the hospital registers the shipment. Also, the employee checks whether the shipment is intended for the hospital. Only after registration, the supplier or forwarder unloads the goods.

Next, the hospital employee retrieves the delivery note from the MaMa system. If the delivery note is not available in digital form, the delivery note enclosed with the goods is removed during unloading. In order to transfer the delivery note information to the MaMa system, the delivery note is recorded either by digitizing the physical delivery note (e.g., recording via scanner), or by uploading the digital delivery note into the system. The system compares the corresponding purchasing data to the delivery note. For this purpose, the delivery note is assigned to a purchasing order using the purchasing order number on each delivery note.

In the case of under-delivery, firstly, the delivered goods are collected, and appropriate information is forwarded to the purchasing unit. The system notes an outstanding quantity and compares it with any subsequent deliveries. Secondly, in the case of an over-delivery, defined deviations are processed normally. If the amount delivered exceeds a specified threshold, the goods are physically sorted out, and appropriate information is passed on to the purchasing unit for clarification. The items are further processed (receipt or return) only after confirmation from the purchasing department. Thirdly, in the case of incorrect or alternative delivery, the items are physically sorted out, and the relevant information is forwarded to the purchasing unit for clarification. In this latter case, the items are not processed until the purchasing department has confirmed acceptance. Otherwise, the items are returned.

The employee identifies all accepted items and captures the item-specific information. The logistic employee captures the UDI attached to the various outer packaging levels using mobile devices. The system verifies the quantity and quality of the delivered items. Relevant item-specific information is the batch number, serial number, and the expiration date, which is retrieved from the UDI. The recording of the item-specific information also enables the comparison of the actually delivered quantities with the corresponding purchasing order and the delivery note data. Finally, the captured data is transferred to the MaMa system for inventory management (posting) collectively with all item-specific information.

For stock items, the employee initiates the storage of the goods after posting. Therefore, the MaMa system assigns each item to the designated storage location. The system creates a warehousing receipt, which is also a signal for the onward physical transfer of the goods. For pass-through items, instead, the MaMa system obtains the storage location, the cost center number, and the cost center description of the requesting department. The system creates a respective hospital-internal delivery note and initializes the shipping to the decentralized warehouses or the place of consumption. Figure 17 summarizes the reference process using VMSN.

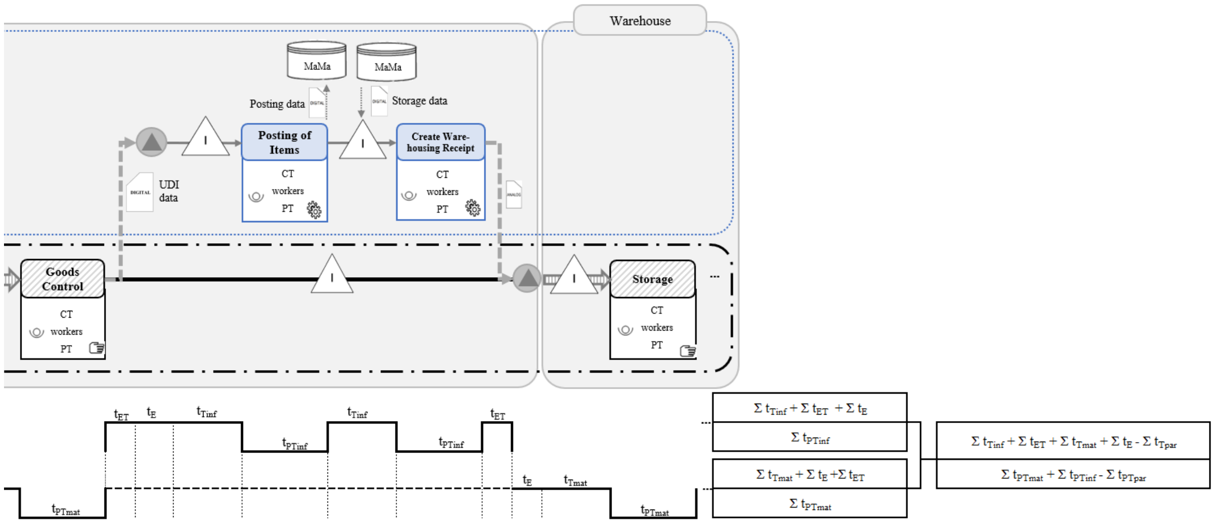


Figure 17 Reference process for goods receipt (extract from goods control to storage)

Goods handling in decentralized storage locations: the reference process covers the transfer, posting, storage, and withdrawal of items. The reference process ends with the demand request. The reference process aims to ensure that medical and care processes are efficiently supplied with medical products (and consumables) with as little effort as possible for nursing staff, while at the same time providing reasonable stock transparency in the storage locations. The reference process is preceded by a demand request, as well as the order (pass-through items) or picking (stock items) and the in-house shipping.

The reference process is based on three premises:

- Supply assistants handle the storage who are responsible for the decentralized storage locations
- Each item has a firmly assigned compartment in the storage location
- Supply assistants primarily pick items (nursing staff only in exceptional cases)

The transport of the goods from the central warehouse or goods receipt to the decentral storage location is carried out either by the hospital's internal delivery traffic or using an automatic transport system. In both cases, the goods are deposited at defined areas in or near the storage locations. The supply assistant receives a message on his mobile device about the provision of the goods at the respective station to keep waiting times as short as possible.

Subsequently, the supply assistant inspects the delivered items analogous to the general goods receipt process at the central warehouse. Here the supply assistant checks whether the delivered goods are intended for the respective ward. Therefore, the supply assistant compares the items to the demand request using the MaMa system, which also points out deviations and outstanding items. The MaMa system also provides an estimate of the expected delivery time for outstanding items. Again, capturing the UDI allows quantity and quality inspection for each item. The inspection includes comparing the batch number with current recalls or examing the expiration dates. For the latter, the system checks whether the average duration of consumption exceeds the expiration date. The system can instruct the supply assistant to consider goods with a short expiration date for early consumption. At last, the system compares the hospital-internal delivery note with the data entered to avoid mistakes.

The posting for inventory management in decentralized storage locations is based on the actual items delivered. The system also memorizes the item-specific data of each good in the respective compartment. Hypothetically, the MaMa system could assign pass-through items to each particular compartment already during the goods receipt process described above. However, this should be avoided to trace the route of the items through the hospital and to ensure data quality.

Following, the items are put into the respective compartments. The system displays the correct compartment to the supply assistant on a mobile device. The items are usually withdrawn by the supply assistant for replenishment at the point of value creation (e.g., treatment, care). If necessary, however, nursing staff can also withdraw goods from the decentralized warehouse when demands are determined ad hoc, for example, based on visual inspection during an emergency.

The withdrawal results in the item being written off from the stock. Actually, every withdrawal would have to be recorded systematically to enable item-based inventory management. However, this is often not practical due to the related high effort. For this reason, the reference process foresees to write off from the stock after a defined quantity was withdrawn (point of reordering). This approach results in an inaccuracy regarding the actual stock at an item level. The stock must, therefore, be checked by the supply assistant using a visual inspection to see whether the point of reordering has been reached. For example, a modular concept with two compartments holds the same number of articles in a row. As soon as all items have been withdrawn from the first compartment, the items are written off. The items in the second compartment are withdrawn until new items are replenished. In a concept with only one compartment, however, the point of reordering (also known as safety stock) must be recognizable. For instance, this could be done using a separator. In any case, the point of reordering covers at least the demand for the duration required for replenishment. When goods are written off, the MaMa system generates a demand request. Demand requests can be collected up to a defined point in time to allow collective purchasing and picking orders. Figure 18 provides an excerpt of the reference process using VMSN.

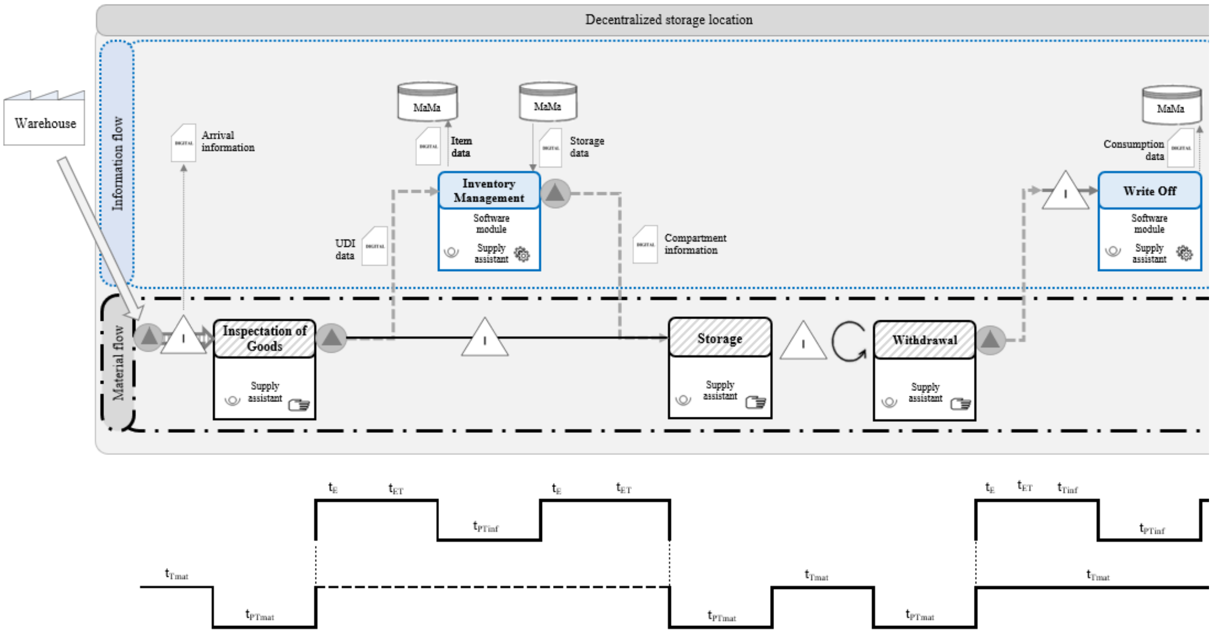


Figure 18: Reference process for goods handling in decentralized storage locations (extract from arrival to write off)

Overall, both reference processes illustrate how to avoid waste in the material logistics of hospitals. Furthermore, they provide a practical approach on how to apply existing digital technologies. The following demonstration example shows how digital technologies were implemented in the context of the research and development project

Demonstration

In order to evaluate the technical feasibility of both reference processes, we developed a prototypical instantiation. The software constitutes a module extension of the MaMa systems in the two participating hospitals (A and B). For this purpose, the relevant data is exported from the MaMa systems. This includes the item master data as well as the purchasing order data. The software module maps both reference processes.

Goods receipt: the software allows digitizing the delivery note of a shipment, in which the physical delivery note is scanned. Alternatively, a digitally existing delivery note is transferred to the demonstrator as a portable document format (PDF) file. The software executes an optical character recognition (OCR) to capture the data on the delivery note. Subsequently, the system obtains the corresponding data of the associated purchasing order from the MaMa system based on the internal order number available on the delivery note. The system compares the delivery note and purchasing order and visually displays any deviations (under-delivery, over-delivery, etc.). Subsequently, the employee records the delivered items by capturing the UDI of each package using a barcode scanner. A glove with an integrated scanner unit is used for this purpose during the pilot phase. This glove avoids additional handling and transmits the recorded data wirelessly to the software module. About two-thirds of the delivered items already possess the UDI or a UDI-like barcode. The system extracts the item identification data from the UDI data and compares it to the identification data available in the item master data (Figure 19). In this way, the software assigns the delivered items to the items on the delivery note and purchasing order. At the same time, item-specific data, such as expiration dates or batch numbers, are automatically captured for all items. Deviations in the number of articles are also displayed visually. Finally, depending on the article type (stock items or pass-through items), the corresponding warehousing receipt or hospital-internal delivery note is created.

Goods handling in a decentralized storage location: the software enables to check the UDI during replenishment. The system compares the item identification data to the item master data of the corresponding compartment. For each compartment, the system notes the expiration dates of each stored item. If the duration of consumption takes longer than the expiration date (based on historical consumption data), the system provides an alert to the supply assistant. This approximate procedure is necessary because complete stock transparency is not achieved during withdrawal. The supply assistant (or nursing staff, if necessary) withdraws the items without digital acquisition. The supply assistant captures the demand only when the defined point of reordering is reached. For this purpose, two approaches have been implemented using radio-frequency identification (RFID).

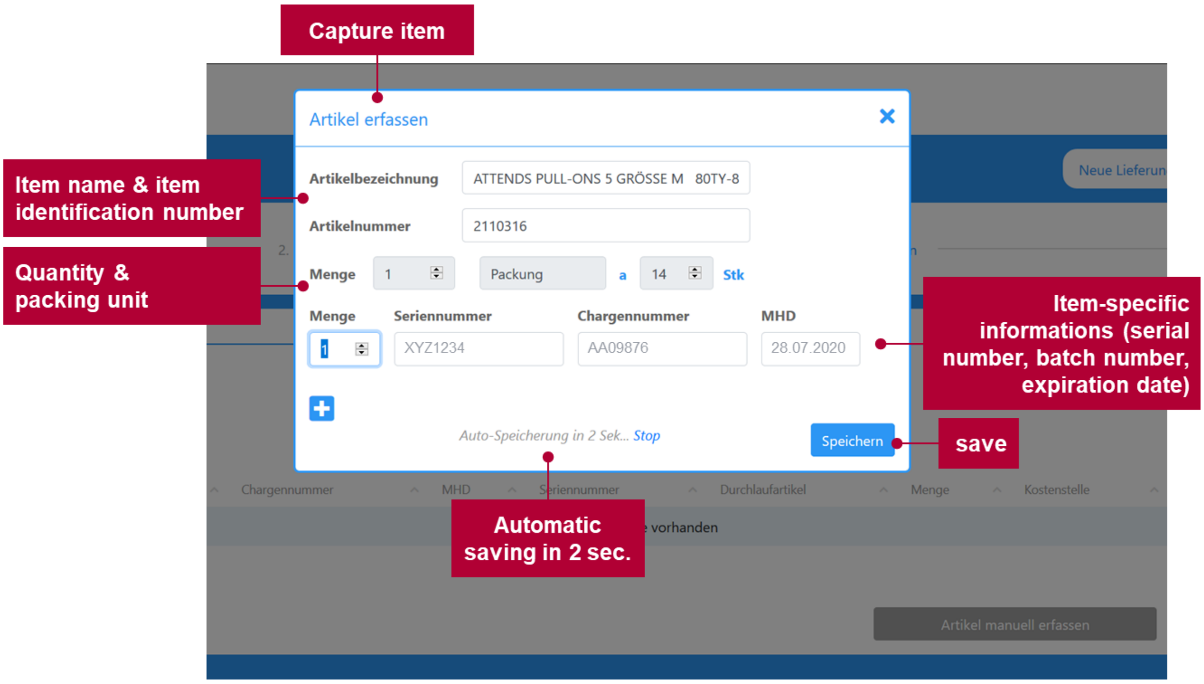


Figure 19: Screenshot of the software module while capturing item-specific data through UDI

In the first approach, the supply assistant holds the corresponding card of the compartment to an RFID-reader. The assistant turns the card to the red side and puts it back to the compartment. Meanwhile, the system captures the item identification number and, thus, the respective demand (see Figure 20). The demand is collected up to defined points in time and then converted into a demand request. Subsequently, the supply assistants The system transfers the demand request to the MaMa system, which triggers all subsequent processes in the central warehouse or purchasing department. Until the system triggers a demand request, a second capturing of item identification number would have no effect. In order to request an increased number of items, the supply assistant can increase the quantity using a tablet in the decentral warehouse. This procedure should prevent multiple accidental requests or the request of the wrong units per item. This first approach is implemented at a hospital in which supply assistants mainly perform withdrawals. The additional effort required for capturing the item identification number saves the effort for sorting the cards upon delivery (see the second approach below) in the further course of the process. Besides, the procedure reduces complexity by only having to turn the cards once (as compared to the current process, see above). In the second approach, the corresponding card is thrown into an RFID-box. This approach has been implemented in an environment where the nursing staff mainly does the removal. The approach reduces complexity by eliminating the need to turn the cards completely. The box recognizes all inserted cards, and the software triggers a demand request at defined times (similar to the first approach). During the storage of the items, the supply assistant must then reattach the cards to the respective compartments. Therefore, two cards are attached to each compartment to keep the effort as low as possible. This enables the supply assistant to carry out a brief visual comparison. In both approaches, the demonstrator provides an overview of outstanding demands that have not been delivered or have only been partially delivered. Manual matching and the maintenance of lists is thus completely unnecessary.

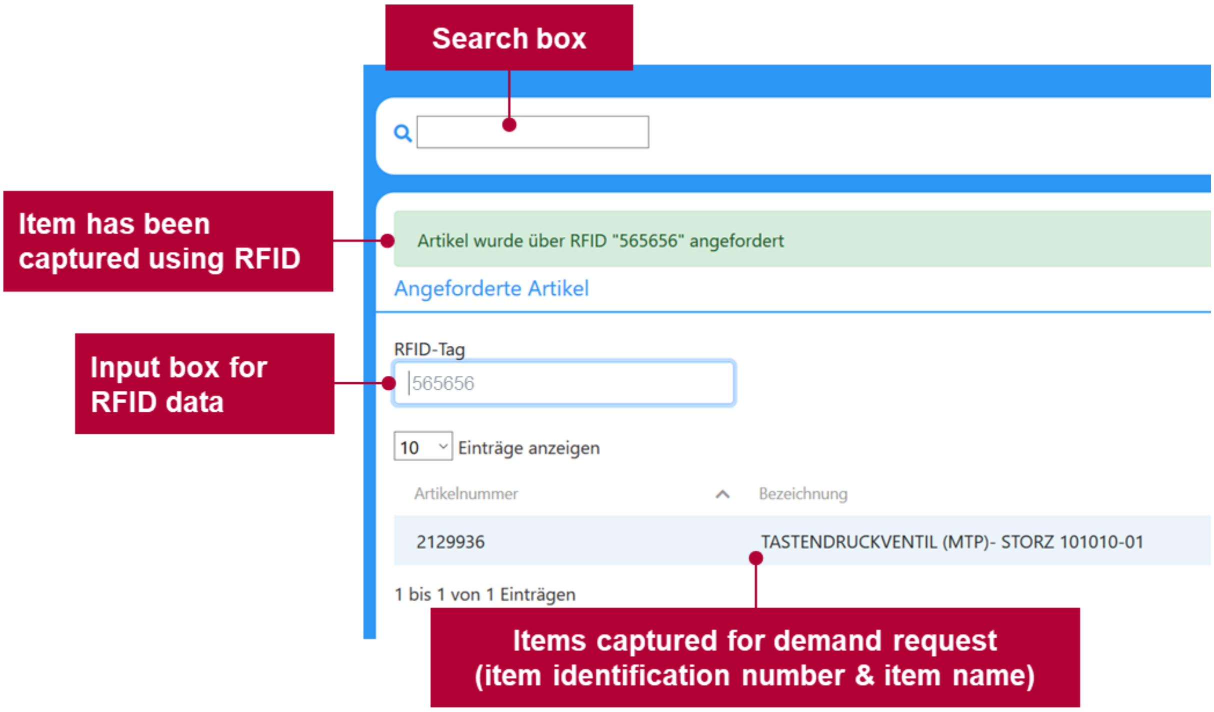


Figure 20: Screenshot of the software module while capturing an actual demand via RFID

To sum up, the implementation of a software module that maps both reference processes shows the technical feasibility. In the next step, the software must be introduced in both hospitals to measure the efficiency gains in a field study. Also, such a field study can allow determining employees' acceptance of the reference processes. Overall, the application of VSMN to design both reference models illustrates the applicability and demonstrates VSMN in practice to develop lean, digitally supported value streams. Besides, the insights from the technical demonstration show that VSMN is suitable to foster the communication between functional departments and IT departments, which develops software artifacts.

Lean Material Logistics at Cardiac Catheter Laboratories using VSMN

In addition to the research and development project, we also utilized VSMN in the scope of another applied-research project to analyze and design material supply processes of the cardiac catheter laboratories at the hospital (M). Together with some employees of the hospital, we conducted a qualitative and quantitative process assessment. The results were visualized using VSMN to examine the value stream. Subsequently, with the help of VSMN, we jointly pre-sented the target state of the process. In the following, we introduce a brief extract of the rec-orded as-is value stream and the corresponding target state. For reasons of confidentiality, we do not report exact numbers regarding lead times.

While assessing the as-is value stream (Figure 21), we found that materials are stored within three warehouses: outside the laboratories in two distinct warehouses (basement, ground floor) as well as inside the laboratories. The material outside of the laboratories was in parts trans-ferred to the laboratories after a daily inventory audit by a ward employee. To do this, the em-ployee creates a list of required materials by carrying out a visual inspection at the laboratories. The employee has to know where the materials are stored (event *demand in laboratories*). The material then remains in the laboratories until it is used up (the resting time was several months to years depending on the material). Materials are consumed during medical procedures. The medical staff has to record which materials were consumed. Nevertheless, the employee carries out a physical inventory at regular intervals and creates a demand request in the MaMa system. Thus, existing information was not used and therefore remained in an information buffer.

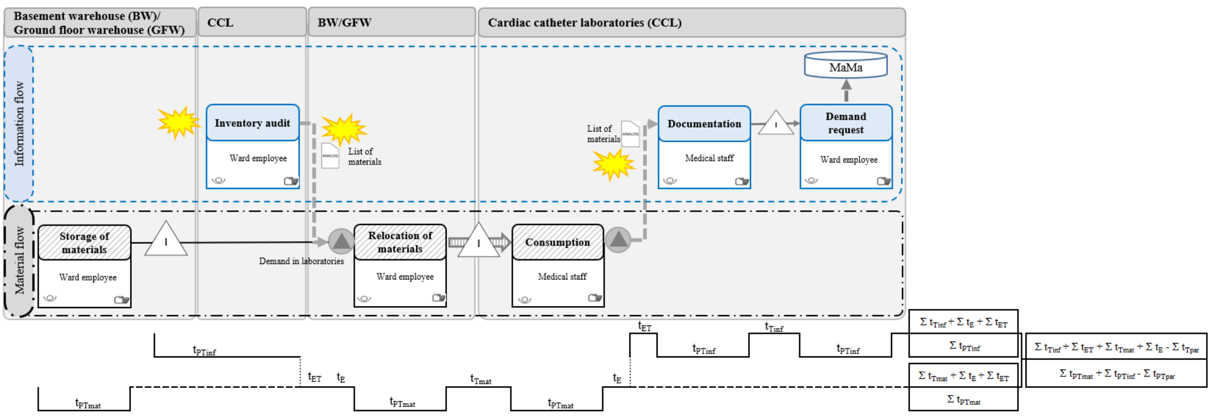


Figure 21: As-is value stream for the material supply process at the cardiac catheter laboratories

Note: kaizen flashes (yellow) indicate waste (Klevers 2009)

We then developed a target state (Figure 22) with the help of semi-structured interviews with employees in purchasing, logistics, and IT of the hospital (M). The target state differs from the actual state in three main points. First, storage and stock transfer are no longer carried out by the ward employees but by a supply assistant. The supply assistant also handles the stock posting and the material checks of the withdrawn articles. These process steps were newly introduced due to the second change, IT-supported inventory management, with the help of a MaMa system. Likewise, an automated inventory posting can also be realized directly from the documentation within the scope of medical treatment. Third, the two warehouses outside the laboratories were combined into one warehouse in a new room, which was created as a systematic withdrawal warehouse. With these changes in process flows and the associated increased use of digital technologies, the hospital believes in achieving a lean value stream in the laboratories. At the time of submitting this thesis, the hospital is in the process of implementing the target state.

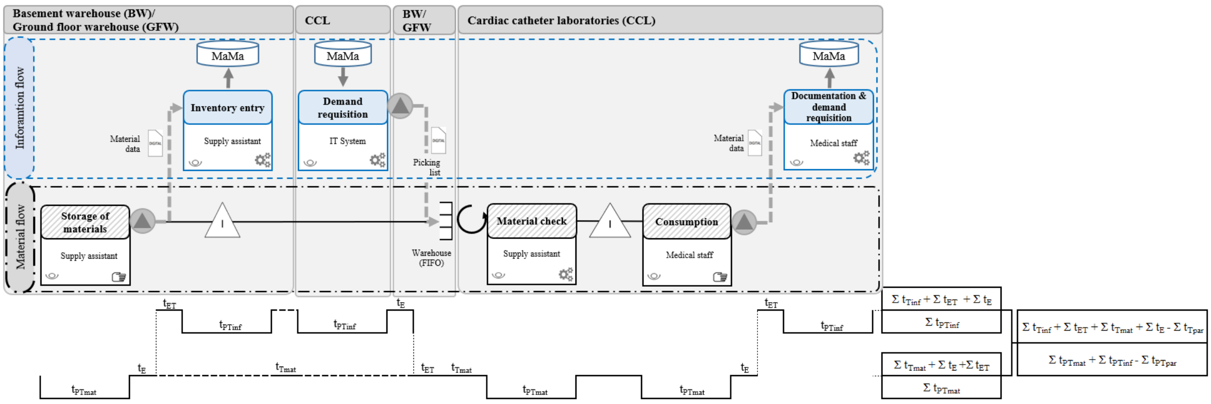


Figure 22: Target state of the value stream for the material supply process at the cardiac catheter laboratories

With the application of VSMN in practice, we ultimately show that our proposed solution also meets the design objectives (introduced in Chapter 5.1) under real-world conditions. We demonstrate that VSMN is suitable for visualizing time dependencies of information flows, analyzing existing value streams, and identifying waste according to inefficient or missing information flows. Furthermore, we demonstrate that VSMN is appropriate for designing value streams and its required information flows and information availability.

Discussion

To sum up, Chapter 5.2 first introduces a reference model for material logistics in hospitals, which arises from a research and development project with two hospitals (A and B). The reference model rests on the analysis of existing material logistics processes. We derive relevant areas of improvement from the insights. Subsequently, we apply VSMN to design lean, digitally supported material logistics processes exemplified by the goods receipt process and the goods handling in a decentralized storage location process. Both processes illustrate the application of digital technologies to achieve efficiency and quality in material logistics based on digitally-enabled transparency. For evaluation purposes, we describe a software solution, which has been developed throughout the research project.

Second, we also outline the results of an applied-research project with another hospital (M). In the course of the project, we applied VSMN to analyze and to design material logistics processes at the cardiac catheter laboratories of the hospital (M). Therefore, we recorded the processes using interviews, observations, and data analysis. Subsequently, we introduced a redesign of the existing value stream to relieve the ward employees and to reduce the lead time of items with the aim to lower stocks.

Together, both applications in practice demonstrate the applicability of VSMN. In Chapter 5.1, we show that VSMN can fill the gap of a missing modeling language for analysis and design of information flows and information available in value streams under consideration of temporal dependencies using expert interviews. However, a shortcoming has been a missing application in practice. We fill this gap with both application scenarios presented in this chapter. However, we do not show how VSMN behaves in practice within other projects and industry sectors. For this purpose, further work may investigate using VSMN in various industries with potential users.

Besides the evaluation of VSMN, the results of this chapter also constitute prescriptive knowledge for material logistics in hospitals from a technocentric perspective. We introduce a reference model that provides design knowledge for the application of digital technologies in goods receipt and goods handling processes. The reference model is derived from a research project with two hospitals and contains knowledge for hospitals (i.e., a class of organizations, see Fettke and Loos 2004). Furthermore, the reference model, including its prototypical implementation, illustrates improvement and, at the same time, a level 1 contribution (instantiation) following Gregor and Hevner (2013). The value stream design for material logistics in the cardiac catheter laboratories at the hospital (M), instead, is not a reference model. However, it also provides level 1 design knowledge as it constitutes an instantiation.

Yet the presented design knowledge is limited to a certain extent. First and foremost, we evaluated our value stream design qualitatively with experts from the hospitals and, in part, developed a prototypical instantiation. However, at the time of writing this thesis, the implementation of the new value stream design is being implemented but is not operational yet. Hence, we do not yet have quantitative data on its efficiency and quality yet. Second, we developed the reference model considering only two hospitals or one hospital for the design of the material logistics processes in the cardiac catheter laboratories. Future research, therefore, should discuss our findings with other hospitals to refine them. Third, our results only describe the design of parts of a hospital's material logistics. Researchers should investigate and bridge the existing gaps, e.g., for picking in the central warehouse. Nonetheless, researchers and practitioners can build on the prescriptive knowledge for material logistics in hospitals and derive more mature design theories in the future.

6. General Discussion and Conclusion

6.1. Summary of Results and Meta-Inferences

One of the essential foundations in IS research is the understanding of IS as socio-technical systems (Briggs et al. 2010; Sarker et al. 2019; Bostrom and Heinen 1977). However, some authors argue that IS scholars in recent years lost this perspective (Sarker et al. 2019). For this reason, Sarker et al. (2019) demand to investigate both sides of the coin: the technical and the social. The authors make three helpful recommendations, which have been published in the leading IS journal *MIS Quarterly*. This doctoral thesis takes up these recommendations as follows.

As a first recommendation, Sarker et al. suggest recognizing "IS problems as consisting of social and technical aspects along a continuum" (Sarker et al. 2019, p. 707). They strive to ensure that IS scholars must develop IS knowledge along the social-technical continuum. For this reason, the doctoral thesis at hand investigates IS from three different perspectives to create IS design knowledge for sustainability and, thus, contribute to sustainable development.

From a balanced, socio-technical perspective, the thesis presents descriptive as well as prescriptive design knowledge. Previous research already took into account the effectiveness of feedback to improve eco-driving behavior. Building on this descriptive knowledge – extended by existing prescriptive knowledge – the thesis develops a nascent design theory (Gregor and Jones 2007) consisting of design principles and an architectural blueprint describing the form and functions of EDFIS. The design theory is based on justificatory knowledge about feedback and norm theory as well as knowledge about (eco-) driving behavior. It builds on previous applications of eco-driving feedback systems and is evaluated utilizing a prototypical instantiation, which is applied throughout a field study. Savings up to 4% on average were achieved. The resulting design artifacts are iteratively refined and finally summarized following the design theory components by Gregor and Jones (2007). In doing so, the design artifact is a level 2 contribution and represents an improvement, which constitutes a more efficient solution compared to previous solutions (Gregor and Hevner 2013). Based on the derived prescriptive knowledge, the doctoral thesis identifies a lack of understanding of how eco-feedback improves driving behavior as research studies to date primarily investigate the effect of eco-feedback on fuel-saving, leaving the changes in driving behavior behind. Therefore, the thesis at hand utilizes the technical possibilities provided by IoT-enabled smart vehicle services. Such services allow access to valuable data sets about driving behavior in real-world settings. Accordingly, the thesis presents a factor model, which is calculated from data of 5.676 drivers making trips over ten weeks. Seven factors describe driving behavior on either operational and strategic levels. On these grounds, eco-feedback has been applied to another 495 drivers, also making trips over ten weeks. After four weeks, they received eco-feedback addressing their driving behavior. Our results indicate that eco-feedback primarily affects acceleration behavior, while speed behavior is interestingly not affected. The acquired descriptive knowledge allows researchers as well as practitioners to build better EDFIS. What is even more, the results reveal how IoT can be adapted to provide new insights on individual driving behavior for research. Therefore, the results are also an exaptation, following Gregor and Hevner (2013), as it is a first approach that extends known solutions to new application domains.

From a rather sociocentric perspective, the thesis at hand presents further primarily descriptive design knowledge on the effect of SNs on behavior in a different context, namely social media. This descriptive knowledge allows deriving implications for the design of social media environments. Social media has been proven to be a breeding ground for the spread of false and intentionally misleading information, commonly known as fake news. Such fake news is a significant threat for societies, not only from a political point of view but also concerning economic and environmental issues. Therefore, the doctoral thesis adds knowledge to an important discussion in the current IS research (e.g., Dennis et al. 2019). Within this discussion, enabling users' capabilities to handle fake news better is a vital countermeasure (Kim and Dennis 2019; Kim et al. 2019). For this reason, the thesis adds to the body of descriptive knowledge about user behavior in social media regarding fake news detection and reporting. Firstly, the thesis investigates an instantiation of related articles within an online experiment simulating a social media environment. As a result, highly controversial additional information is found to be a valuable countermeasure to improve fake news detection rates. Participants within the experiment were able to detect fake news better when exposed to controversial related articles. Interestingly, providing solely accurate additional information did not lead to significantly better detection rates. In contrast, providing additional information in the form of related articles did not improve nor impair the detection of truthful news anyway. Secondly, the thesis conducts two further online experiments in order to derive knowledge of users' reporting behavior. Social psychologists already demonstrated the suitability of SNs in the course of other application domains (e.g., see Chapter 3.1). For this reason, the thesis adapted both injunctive and descriptive SN messages and investigated effects regarding the reporting behavior of fake news in social media. Results indicate that a combination of injunctive and descriptive SNs achieve the best results, which is in line with the seminal works of Cialdini et al. (1990) and Goldstein et al. (2008) on the focus theory of normative conduct. While solely injunctive SNs also showed effects, the application of solely descriptive SNs – neither being mainly positive nor negative – showed any effects. Overall, the findings on fake news detection, as well as reporting behavior, add descriptive knowledge to current debates in IS research from a sociocentric perspective. Both online experiments adapt approaches from social psychology for social media platforms and, thus, investigate user behavior in digital environments. The results, therefore, extend the understanding of existing kernel theories and provide implications for the design of social media platforms. In doing so, both chapters also describe level 1 artifacts and pose an exaptation following Gregor and Hevner (2013).

From a rather technocentric perspective, the thesis at hand presents primarily prescriptive knowledge for the digital transformation of value creation processes in order to avoid waste due to unavailable or insufficient information. Such waste is not only resource overuse and increasing risk of unused and thus expired materials, but also inadequate utilization of capacities. Firstly, the thesis presents a new modeling language, namely the *Value Stream Modeling and Notation (VSMN)*, which supports modelers and model users to analyze their value streams regarding temporal dependencies due to information availability. Subsequently, VSMN allows designing and visualizing target states of lean, digitally supported value streams. Throughout Chapter 3 and Chapter 4, the thesis utilizes information visualization to improve user behavior individually (driving behavior) and collectively (report and detect fake news). VSMN extends this and employs information visualization to support individual analysis and designs in organizations but also to improve collective discussions of teams to achieve lean, digitally supported value streams. VSMN is a design artifact and constitutes a model following Hevner et al. (2004). VSMN, therefore, is a level 2 contribution, being a nascent design theory (Gregor and Hevner 2013). As existing approaches already allow to analyze and design value streams but illustrate insufficient approaches regarding the novel possibilities of digital technologies, VSMN also constitutes an improvement following Gregor and Hevner (2013). Secondly, the thesis presents a reference model for material logistics processes, which originates from the research and development project using VSMN. Furthermore, the thesis provides insights into the analysis and design of material logistics in cardiac catheter laboratories using VSMN, too. Various studies already examined solutions to improve material logistics and supply chains of hospitals. The value stream designs extend existing solutions and provide an additional viewpoint concerning the application of digital technologies in hospital material logistic processes. Promising insights from the hospitals indicate that rising challenges, such as increasing patient numbers and cost pressure, are diminished by the proper use of digital technologies. Overall, the findings on VSMN and the design of lean, digitally supported material logistic processes in hospitals add prescriptive knowledge to current research debates at the interface to lean management (e.g., Bertagnolli 2018) and hospital logistics (e.g., Volland et al. 2017). The results arise mainly from qualitative research methods such as expert interviews and a case study. Such methods are common for naturalistic evaluations (Sonnenberg and Vom Brocke 2012; Venable et al. 2012) and, thus, the evaluation of rather socio-technical artifacts (Venable et al. 2012). The results, therefore, extend the body of prescriptive knowledge and enables researchers and practitioners to transform value creation processes effectively. Moreover, the results also provide a reference model for lean, digitally supported hospital logistic processes. In doing so, both results pose an improvement following Gregor and Hevner (2013) from a rather technocentric perspective.

Taking up the three perspectives (social-technical, sociocentric, and technocentric), the doctoral thesis at hand contributes design knowledge to both – the descriptive and prescriptive – knowledge bases in IS research. While the thesis takes first a socio-technical perspective with regard to a balanced interplay of humans and technology to achieve energy-efficient behavior, the other perspectives concentrate on one side of the coin at a time, while not ignoring the other side. The sociocentric perspective mainly investigates users' behavior and problem-solving capabilities in digital environments. The technocentric perspective derives solutions primarily focused on the application of digital technologies and ensures applicability and acceptance at the users' side at the same time.

Taking perspectives along the continuum is in line with Sarker et al.'s second recommendation to accept "variations of social-technical relationships" (Sarker et al. 2019, p. 709). Sarker et al. (2019) claim to exploit the various opportunities of the social and technical interplay to apply different approaches to investigate phenomena of interest (Sarker et al. 2019). Flexible examining the social and the technical allows taking up problems from other disciplines and to address relevant issues employing IS research. According to the above summary, the doctoral thesis contributes to achieving this recommendation. It provides an IS perspective on different phenomena. At the same time, the IS discipline is brought together with other disciplines (behavioral science, social psychology, and lean management). The thesis thus helps to expand the stakeholder base of IS (Sarker et al. 2019). With their second recommendation, Sarker et al. (2019) imply two effects.

Firstly, they derive that redirecting IS research and concentrating more on the socio-technical interplay forces IS scholars to think about appropriate methodological approaches. The presented thesis uses pluralistic methodological approaches. While the socio-technical perspective and the sociocentric perspective primarily focus on quantitative methods, the later technocentric perspective applies rather qualitative methods to claim evidence. Qualitative and quantitative methods are conventional evaluation methods in design science research (Sonnenberg and Vom Brocke 2012; Venkatesh et al. 2013; Venable et al. 2016). However, they also bring along another vital advantage. Following Venkatesh et al. (2013), multi-method approaches allow an extension of the existing knowledge from different perspectives. With the help of quantitative data from a field study as well as from an IoT-based smart vehicle service, the findings from a socio-technical perspective add to a clearer understanding of eco-driving feedback systems.

Analyzing data from IoT-based services, for instance, allows investigating human behavior with technology in real-world applications and, therefore, leaves biases (e.g., the Hawthorne effect, see Rice 1982) behind. Similarly, the findings from a sociocentric perspective arise from online experiments, which aim at simulating a social media environment. The results contribute to current debates in IS research and thus expand the existing understanding of kernel theories in a different application domain but also constitute level 1 artifacts to improve social media users' capabilities regarding false information. To date, for instance, researchers applied field studies (e.g., Cialdini et al. 1990; Cialdini 2003; Goldstein et al. 2008), laboratory experiments (e.g., Moravec et al. 2019), or surveys (e.g., Kim and Dennis 2019) to collect data. Online experiments applied in this thesis, expand the set of methods to research from a rather sociocentric perspective. In contrast, the findings from a technocentric perspective arise from qualitative methods, such as expert interviews or case studies. The results yield useful solution artifacts for the digital transformation of value creation processes altogether and material logistics processes in hospitals in particular. Even more important, both artifacts are developed and evaluated using methods, which also take a user-centric perspective into account. That is, besides its effectiveness, VSMN has been assessed regarding its acceptance and applicability through members of potential target groups through expert interviews. Thus, the resulting artifacts but also the applied methods add to a prescriptive knowledge base from divergent views. The artifacts – primarily VSMN – allow a novel consideration of value streams in times of digitalization. Expert interviews and case studies allow evaluating different quality attributes, such as effectiveness, usability, and acceptance. In summary, the doctoral thesis indicates how different methods in IS research can be used to study socio-technical systems.

Secondly, Sarker et al. derive another implication from their second recommendation. They argue that refocusing on socio-technical perspectives allows engaging with information as a central construct of IS discipline (Sarker et al. 2019). Following Lee et al. (2015), IS artifacts are the fundamental concept of IS research, which consists of technology artifacts, information artifacts, and social artifacts. As the results of the doctoral thesis constitute design knowledge for IS artifacts from different perspectives, it also investigates information artifacts as the central construct. More precisely, EDFIS constitute an interplay of technology artifacts (the feedback system which measures behavior and provides the feedback information, e.g., mobile devices), information artifacts (the feedback information itself, e.g., an eco-score), and social artifacts (the drivers who exhibit driving behavior and adjust their behavior according to the feedback information). The prescriptive knowledge describes how to design the interplay of all three artifacts. The descriptive knowledge contributes to a better understanding of the social artifacts (i.e., the drivers and their driving behavior) and their behavioral improvements in dependency on the information artifact (i.e., the improvements according to feedback information). Similarly, the interplay of information artifacts and social artifacts also plays a central role in improving users' capabilities to handle fake news in social media environments. Here, the social media platform, as the technology artifact, is primarily the application domain. More interestingly, additional information artifacts (the SN messages and related articles) have an influence on the social artifact (being the users, who better detect or more often report fake news). In particular, the descriptive SN message (being an information artifact) makes the behavior of the other social media users transparent (the temporal relationship between the individuals who are in touch with a fake news story constitutes a social artifact). Instead, VSMN itself is an information artifact, which allows to communicate complex circumstances, such as reference models for material logistics in hospitals (which itself consists the interplay of all three artifacts) between stakeholder (who constitute the social artifact). Thus, the research within this thesis is along the social-technical continuum and, moreover, also engages with information artifacts as the central component of the research objectives.

As a third and last recommendation, Sarker et al. suggest connecting "humanistic and instrumental outcomes in a synergistic manner" (Sarker et al. 2019, p. 710). Besides the fact that IS research in recent years predominantly investigated the social in a technical context, Sarker et al. also criticize that IS research primarily focuses on instrumental goals rather than humanistic goals. Instrumental goals strive for instrumental outcomes, such as increasing productivity. Humanistic goals instead aim at achieving humanistic outcomes, such as increasing job satisfaction (Wallace et al. 2004; Sarker et al. 2019). The socio-technical theory suggests that there is a virtuous cycle in IS research, which connects both goals. However, IS scholars should develop solutions that address instrumental and humanistic goals equally, as increasing humanistic outcomes do have a positive effect on instrumental outcomes. This synergy allows researchers to achieve long-term objectives (Sarker et al. 2019). The doctoral thesis at hand argues that this is also the case regarding sustainable development. A lifestyle in balance with nature is essential for the long-term well-being of humans. Therefore, the presented findings, to some extent, achieve instrumental and humanistic goals at the same time.

Clearly, the presented domain-specific modeling language VSMN mainly aims at achieving instrumental goals. Avoiding waste that results from insufficient information availability results in economic advantages such as higher productivity, lower stocks, shorter lead times, and, therefore, fewer costs. In the same vein, the reference model for lean, digitally supported material logistics in hospitals contributes to the same instrumental goals. As a consequence, avoiding waste and thus an overuse of resources as well as garbage (i.e., waste resulting due to unused materials) contributes to achieving *sustainable economies*. However, sustainable economies also address humanistic goals in the long-term (e.g., less poverty, sustainable economic growth, avoiding resource overuse). Another humanistic goal, closely related to lean hospital logistic processes, is affordable healthcare costs. That is, logistic processes in hospitals are primary cost drivers in hospitals (Ross and Jayaraman 2009). Reducing efforts within those processes results in financial advantages. No less important, appropriate logistic processes reduce the involvement of nursing staff in logistic tasks and relieve the burden to have more capacity for the care of patients. Both advantages serve *human well-being*.

In contrast, achieving a more sustainable lifestyle through more energy-efficient behavior contributes to humanistic goals. *Decarbonization* is essential to mitigate climate change, which in turn is a significant threat to our environment, entailing the risk of dangerous diseases, natural disasters, or economic restrictions (i.e., unemployment, poverty, etc.). For this reason, energy-efficient behavior also contributes to *human well-being*. The presented design knowledge about EDFIS helps to achieve those goals even if they can only have a small effect. Nevertheless, energy-efficient behavior is essential against the background of increasing energy demand (Messerli et al. 2019). *Decarbonization* addresses primarily humanistic goals but also instrumental goals. For instance, EDFIS allow car manufacturers and their suppliers to implement better onboard IS and thus offer more environmentally friendly products. Similarly, providers of smart vehicle services also may add eco-driving feedback functionalities to their services to reach environmentally friendly customer segments. Last but not least, companies and organizations, which have employees in the field, benefit from energy-efficient driving behaviors of their employees that lead to decreasing fuel costs.

Finally, designing social media platforms to improve users' capabilities to detect and report fake news adds to achieve both instrumental and humanistic goals. Firstly, false and misleading information is often spread to influence economies and particularly financial markets, for instance, to affect stock values, but also to increase revenue from news consumption through advertising (e.g., Maasberg et al. 2018). Thus, avoiding fake news contributes to *just economies* and thus helps to achieve instrumental goals. Secondly, in recent times fake news was primarily spread to influence humans' believes and distort political discourses (Allcott and Gentzkow 2017). The spread of intentionally false information also causes populism. Another threat of fake news is the so-called infodemics, which describe the dissemination of incorrect information about virus pandemics as the recent Covid-19 outbreak (Cinelli et al. 2020). False information can harmfully influence people's behavior and amend the effects of countermeasures. Such infodemics can have serious health consequences (Cinelli et al. 2020). Moreover, false information is often published to deny the environmental threat of climate change (van der Linden et al. 2017). Therefore, designing effective countermeasures to avoid the spread of false information contributes to rather humanistic goals. For instance, honest and fact-based journalism on pandemics (e.g., Covid-19) is essential to ensure health and, thus, *human well-being*. Similarly, news about climate change and sustainable development are necessary to anchor the environmental threat in people's minds and thus achieve sustainable development and *decarbonization*.

Overall, the doctoral thesis at hand addresses both humanistic and instrumental, goals and adds knowledge to the lever science and technology to support essential entry points for transformation (Messerli et al. 2019), namely *human well-being and capabilities, sustainable and just economies*, and *decarbonization*. The thesis extends existing design knowledge to achieve sustainable development and sustainability from an IS research perspective. Therefore, the thesis takes three perspectives along the social-technical continuum. In this way, it also contributes to the demand for reflection on the original orientation of the IS discipline.

6.2. Limitations and Outlook for Future Research

Like all research, the doctoral thesis at hand suffers from limitations. However, these limitations are also starting points for future research. In the following, a summary of the limitations and an outlook for future research are given.

First, regarding the results from a socio-technical perspective, in Chapter 3.1, the results are not complete or universal. Foremost, the results do not explain which feedback works best for a specific individual. The driver receives the appropriate feedback and ignores the other types when combining different types of feedback. Besides, driving behavior depends on the personality (e.g., Shahab et al. 2013), and this also applies to the effect of feedback (Ilgen et al. 1979; Brouwer et al. 2015). Thus, future research should further investigate the dependency of different feedback types and personality traits and derive implications refining descriptive knowledge. Second, the thesis presents only the effects of anonymous social comparison. Yet research suggests that comparison with known persons can have other impacts (e.g., Schultz et al. 2007; Foster et al. 2010). Third, there are limitations in the field study design. The participants used the mobile EDFIS instantiation for only eight weeks. Further, there was no control group, not obtaining feedback. It might be the case that the participants were explicitly eager to change their driving behavior (self-selection bias), who then did so more or less independently of the EDFIS specific functionality. Further research should verify the effectiveness of design artifacts in different settings and with different instantiations. Besides, the design artifacts have not been considered concerning safety-critical factors, which may influence the design of a system (Jamson et al. 2015). Finally, different trip profiles like city streets, highways, speedways, or various covered distances may influence the results as several authors illustrate (e.g., Boriboonsomsin et al. 2010). The thesis makes these shortcomings the subject for future research.

In Chapter 3.2, the real-world field study has some restrictions. The mobile app only provided eco-feedback, and it is not clear whether all participants regularly checked their eco-feedback or indeed received the push notifications. Furthermore, it is not clear whether the cars involved in the study were driven only by participants. Thus, the presumed effects of eco-feedback on driving behavior might be more robust when the feedback would be more salient in the car. Future work could ensure user participation and focus on specific groups of drivers, selected based on either similar driving behavior or personal factors. Focusing will allow further investigations into the effects of eco-feedback on specific sub-groups and will, thus, enable more customized and useful feedback in a real-world setting. Furthermore, the data set is limited by the variables that the smart vehicle service provider measured and disclosed for the study. As a consequence, the factor model does lack certain variables, like gear-shifting, type of road, or the actual fuel consumption. Additional variables could enhance the factor model and further investigate and improve the effectiveness of feedback. Hence, the factor model could be extended by measuring other variables to describe driving behavior in more detail. Based on the factor model, more sophisticated analyses are conceivable, which could consider, for example, speed or RPM as no linear functions in terms of fuel consumption and the effectiveness of feedback. Moreover, the sample was restricted to customers of the smart vehicle service, which implies a limitation to Germany and possibly a self-selection bias as customers are presumably more interested in vehicles and potentially care about their driving style. The study only considered the effect of eco-feedback in the short term as the data set only contains information about the variables for ten weeks in total. Future research, therefore, could gain insights from other samples and investigate driving patterns to evaluate different types of feedback to increase impact, as the feedback applied presumably influenced participants with environmental awareness. Finally, the analysis of the effect of eco-feedback does not include a control group and, thus, might be affected by unmeasured or uncontrolled external conditions, i.e., changes in weather between the period prior and after the launch of the eco-feedback. As a consequence, future research might collect additional data that describe moderating effects, which can affect driving behavior.

Second, regarding the results from a sociocentric perspective, the online experiment to investigate fake news detection behavior in social media (Chapter 4.2) is also restricted due to some limitations and provides starting points for future research. Firstly, the generalizability of the study needs to be validated in future research because the age range of participants was limited to relatively young participants being Germans. Further tests with different age groups and nationalities should be carried out. Further, future research should investigate the various factors crucial to the detection of fake news (Flintham et al. 2018). Previous research has shown that the source of an article is an essential factor for its perceived credibility (Amin et al. 2009). The participants' previous knowledge of the topics and prior believes covered in the articles may have an influence on detection behavior and should also be a focus of future research. Also, not all selected related articles are of the same relevance to the main article. Future research could investigate our results with another sample of articles different from those in the study at hand. Furthermore, the study focuses on short-term behavior. Future attempts should investigate long-term behavior to provide a more holistic understanding of behavioral patterns. In this sense, it will also be essential to examine users' reactions to the related articles disclosure nudge once it has been in use for an extended period and is no longer a novel feature. Lastly, the study took place in an experimental environment, a fact that may have influenced the behavior of the participants. Future research could gather data on real social media usage to address this issue.

Subsequently, the online experiment to investigate fake news reporting behavior in social media (Chapter 4.3) provides starting points for future research from the following restrictions. The generalizability of the empirical results is limited by the design of the experimental materials (inspired by Facebook, one wording for the injunctive SN, one way of displaying the descriptive SN) and by our participants (young, German, educated, active social media users). Furthermore, in practice, users of social media platforms cannot always be asked to confirm the same injunctive SN. A variation of the wording may help to confront the user more often with injunctive SNs and to promote desirable behavior. The identified motivational factors and hurdles for reporting fake news may support the development of injunctive SN messages. Future research might empirically test the derived hypotheses with more diverse or complimentary samples in other forms of social media that use other design elements or have a different structure. In this context, future work should also deal with the boundary conditions of the descriptive SN in more detail. Going beyond artificial experimental settings to field experiments on live social media platforms would significantly strengthen external validity. Next, future work should explore whether SNs from the direct social environment have a stronger influence and how SNs and factors affecting cognition and judgment. SN-based interventions are appropriate socio-technical tools to foster active user engagement, which allows for follow-up measures such as user source rating or third-party fact-checking. We suspect that descriptive SNs also improve the ability to detect fake news. Future research may investigate this effect. Another limitation occurs from the headlines used throughout the experiments. The composition of the posts' headlines, the posts' images, the perceived reputation of the posts' sources, and subjects, as well as the users' familiarity with the news, might all affect the identification of fake news. The experiment did not explore these factors related to detection but solely focused on reporting behavior. Therefore, future research should confirm the reported results using other headlines and news topics.

Third, regarding the results from a technocentric perspective, the design knowledge to foster the digital transformation of value creation processes also provides starting points for future research. Throughout Chapter 5.1, the thesis presents VSMN, which mostly suffers limitations from the iterative approach taken. The research design cannot formally provide evidence that all possible scenarios and constructs of VSMN have been covered. However, because of the positive evaluation by five experts from different companies, VSMN likely includes the most vital constructs and scenarios. The evaluation ensured that VSMN addresses an urgent problem and represents a meaningful solution and is understandable and applicable. From a methodological perspective, a modeling language is presented. Importantly, VSMN is not a modeling method. Existing methods, as the VSM (Klevers 2009) and its extensions (Hartmann et al. 2018; Busert and Fay 2019), must be questioned against the background of VSMN and further developed if necessary. At last, the development of VSMN is based on the notation of VSM. However, also other modeling languages, such as BPMN, for instance, could be promising starting points leading to valid design solutions, too.

In Chapter 5.2, the thesis presents results from two projects at different hospitals to evaluate VSMN while it is applied in practice. Nevertheless, future work will have to validate and further develop VSMN in further applications, projects, and industry sectors, such as the automotive sector or the commerce sector. Furthermore, VSMN must be applied by other potential users to improve its acceptance. Besides, Chapter 5.2 develops prescriptive knowledge for material logistics in hospitals. However, the results suffer limitations as the prototypical instantiation has not yet been deployed. Hence, the chapter only demonstrates technical feasibility. Future research should show its applicability and efficiency in practice. Furthermore, the results originate from two hospitals or only one hospital in the cardiac catheter laboratory case. Researchers, therefore, should refine the results using insights from various other hospitals. At last, the results only describe parts of material logistics in hospitals. Thus, future research should add knowledge to a brighter design.

From an overarching viewpoint, the multitude of perspectives and methodologies used and integrated throughout the thesis corroborate each other. At the same time, the combination of perspectives and methodologies is not complete in the sense that, for example, a quantitative evaluation of VSMN (e.g., data from a field study applying the results of a value stream design) is missing. Given that each methodology has its own strength and weaknesses, by far, not every methodology is suitable for every research question. For each specific perspective, research question, or design objective and given the state of knowledge, the thesis aimed to apply the most appropriate research methodology. However, further complementary research (e.g., the application of SN on social media platforms to collect data in a real-world setting) would further strengthen the results. Also, the thesis presented descriptive and prescriptive knowledge, including, for example, multiple level 1 design artifacts and level 2 nascent design theories. With future research, they may mature to more abstract and widely applicable level 3 mid-range (or eventually even grand) design theories.

Overall, and beside the mentioned limitations, the thesis adds valuable design knowledge to foster sustainable development. Furthermore, the results indicate that appropriate and goal-oriented design is suitable to achieve sustainable development regarding the different dimensions (environmental, economic, social) and addresses essential entry points to overcome one of the most significant challenges of humanity and the greatest challenge of the 21st century. However, more research is needed. Besides the starting points mentioned, IS research should increase research activities to elaborate on the promising disruptive potential of IS to achieve sustainability. This claim is in line with various research agendas and calls for action over the last decade. Starting a decade ago Melville (2010) and Watson et al. (2010) but also recent contributions of Gholami et al. (2016), Seidel et al. (2017), and Henkel and Kranz (2018) provide convincing arguments that still a lot of work is needed in IS research. Future research should consequently publish valuable insights and pursue the discussion within the research strand *IS for Environmental Sustainability.*

As another example, *Information and Communication Technology for Development* (ICT4D) is a research strand that has been gaining in importance again in recent years (Walsham 2017). ICT4D is “concerned with the use of ICTs for international development” (Walsham 2017, p.

18). ICT4D research is increasingly focusing on developing countries and often addressing the weak members of societies (Walsham 2017). Future research in many disciplines should involve themselves in investigating the role of ICTs for development (Walsham 2017).

Together, bot research strands can make significant contributions to sustainable development and our future, as a consequence. However, today we fail to pursue sustainable development and the SDG at the necessary speed. Science and technology must further elaborate pathways to address entry points and develop new solutions as a lever for transformation to achieve sustainable development. This elaboration requires not only research from an IS perspective but also the active engagement of various research disciplines. The doctoral thesis at hand, therefore, concludes with the statement of United Nations' secretary-general António Guterres, who postulates science as “our great ally in the efforts” (Messerli et al. 2019, p. XI). We still have a long way to go.

References

Agresti A, Kateri M (2011) Categorical Data Analysis. In: Lovric M (ed) International Encyclopedia of Statistical Science. Springer, Berlin, Heidelberg

Allcott H (2011) Social norms and energy conservation. Journal of Public Economics 95(9-10), pp 1082 - 1095

Allcott H, Gentzkow M (2017) Social Media and Fake News in the 2016 Election. National Bureau of Economic Research, Cambridge, MA

Allison CK, Stanton NA (2019) Eco-driving: the role of feedback in reducing emissions from everyday driving behaviours. Theoretical Issues in Ergonomics Science 20(2), pp 85 - 104

Allweyer T (2015) BPMN 2.0 - Business Process Model and Notation: Einführung in den Standard für die Geschäftsprozessmodellierung, 3., aktualisierte und erweiterte Auflage. BOD - Books on Demand, Norderstedt

Alter S (2013) Work System Theory: Overview of Core Concepts, Extensions, and Challenges for the Future. Journal of the Association for Information Systems 14(2), pp 72 - 121

Amin A, Zhang J, Cramer H, Hardman L, Evers V (2009) The effects of source credibility ratings in a cultural heritage information aggregator. In: Tanaka K (ed) Proceedings of the 3rd Workshop on Information Credibility on the Web. ACM, New York, NY, pp 35 - 42

Andor MA, Gerster A, Gillingham KT, Horvath M (2020) Running a car costs much more than people think - stalling the uptake of green travel. Nature 580, pp 453 - 455

Bagozzi RP, Yi Y (1988) On the evaluation of structural equation models. Journal of the Academy of Marketing Science 16(1), pp 74 - 94

Balmas M (2012) When Fake News Becomes Real. Communication Research 41(3), pp 430 - 454

Barron G, Yechiam E (2002) Private E-Mail Requests and The Diffusion of Responsibility. Computers in Human Behavior 18(5), pp 507 - 520

Barthel M, Mitchell A, Holcomb J (2016) Many Americans Believe Fake News Is Sowing Confusion. http://www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/. Accessed 21 February 2020

Bätz A, Gimpel H, Heger S, Wöhl M (2020) Driving Sustainably – The Influence of IoT-based Eco-Feedback on Driving Behavior. In: Bui T (ed) Proceedings of the 53rd Hawaii

International Conference on System Sciences. Hawaii International Conference on System Sciences

Baumgartner SE, Valkenburg PM, Peter J (2011) The Influence of Descriptive and Injunctive Peer Norms on Adolescents’ Risky Sexual Online Behavior. Cyberpsychology, Behavior and Social Networking 14(12), pp 753 - 758

Bergquist M (2020) Most People Think They Are More Pro-Environmental than Others: A Demonstration of the Better-than-Average Effect in Perceived Pro-Environmental Behavioral Engagement. Basic and Applied Social Psychology 42(1), pp 50 - 61

Bernard J-G, Dennis AR, Galletta DF, Khan A, Webster J (2019) The Tangled Web: Studying Online Fake News. In: Krcmar H, Fedorowicz J, Boh WF, Leimeister JM, Wattal S (eds) Proceedings of the 40th International Conference on Information Systems, ICIS 2019. Association for Information Systems

Bernedo M, Ferraro PJ, Price M (2014) The Persistent Impacts of Norm-Based Messaging and Their Implications for Water Conservation. Journal of Consumer Policy 37(3), pp 437 - 452

Bertagnolli F (2018) Lean Management: Einführung und Vertiefung in die japanische Management-Philosophie. Springer Gabler, Wiesbaden

Blair CA, Foster Thompson L, Wuensch KL (2005) Electronic Helping Behavior: The Virtual Presence of Others Makes a Difference. Basic and Applied Social Psychology 27(2), pp 171 - 178

Boriboonsomsin K, Vu A, Barth M (2010) Eco-Driving: Pilot Evaluation of Driving Behavior Changes Among U.S. Drivers, Riverside: University of California Transportation Center

Bostrom RP, Heinen JS (1977) MIS Problems and Failures: A Socio-Technical Perspective. Part I: The Causes. MIS Quarterly 1(3), pp 17 - 32

Bouhdadi M, Balouki Y, Chabbar E (eds) (2007) Meta-Modelling Syntax and Semantics of Structural Concepts for Open Networked Enterprises. Computational Science and Its Applications – ICCSA 2007. Springer Berlin Heidelberg

Brandon G, Lewis A (1999) Reducing household energy consumption: a qualitative and quantitative field study. Journal of Environmental Psychology 19(1), pp 75 - 85

Brewer GD, Stern PC (2005) Decision making for the environment: Social and behavioral science research priorities. National Academies Press Washington, DC

Briggs RO, Nunamaker JF, Sprague RH (2010) Special Section: Social Aspects of Sociotechnical Systems. Journal of Management Information Systems 27(1), pp 13 - 16

Bronstein MV, Pennycook G, Bear A, Rand DG, Cannon TD (2019) Belief in Fake News is Associated with Delusionality, Dogmatism, Religious Fundamentalism, and Reduced Analytic Thinking. Journal of Applied Research in Memory and Cognition 8(1), pp 108 - 117

Brouwer RFT, Stuiver A, Hof T, Kroon L, Pauwelussen J, Holleman B (2015) Personalised feedback and eco-driving: An explorative study. Transportation Research Part C: Emerging Technologies 58, pp 760 - 771

Busert T, Fay A (2019) Extended Value Stream Mapping Method for Information based Improvement of Production Logistics Processes. IEEE Engineering Management Review, pp 119 - 127

Caraban A, Karapanos E, Gonçalves D, Campos P (2019) 23 Ways to Nudge. In: Brewster S, Fitzpatrick G, Cox A, Kostakos V (eds) Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI ‘19. ACM Press, New York, New York, USA

Chatfield AT, Reddick CG, Choi KP (2017) Online Media Use of False News to Frame the 2016 Trump Presidential Campaign. In: Hinnant CC (ed) Proceedings of the 18th Annual International Conference on Digital Government Research. ACM, New York, NY, pp 213 - 222

Chaykowski K (2017) Facebook Expands Fight Against Fake News With Automatic, Related Articles. https://www.forbes.com/sites/kathleenchaykowski/2017/08/03/facebook-expands-fight-against-fake-news-with-automatic-related-articles/#59fa46b31d2c. Accessed 9 June 2019

Cialdini RB (2003) Crafting Normative Messages to Protect The Environment. Current Directions in Psychological Science 12(4), pp 105 - 109

Cialdini RB (2007) Descriptive Social Norms as Underappreciated Sources of Social Control. Psychometrika 72(2), pp 263 - 268

Cialdini RB, Demaine LJ, Sagarin BJ, Barrett DW, Rhoads K, Winter PL (2006) Managing Social Norms for Persuasive Impact. Social Influence 1(1), pp 3 - 15

Cialdini RB, Kallgren CA, Reno RR (1991) A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior. In: Advances in Experimental Social Psychology, vol 24. Elsevier, pp 201 - 234

Cialdini RB, Reno RR, Kallgren CA (1990) A Focus Theory of Normative Conduct: Recycling The Concept of Norms to Reduce Littering in Public Places. Journal of Personality and Social Psychology 58(6), pp 1015 - 1026

Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, Zola P, Zollo F, Scala A (2020) The COVID-19 Social Media Infodemic. http://arxiv.org/pdf/2003.05004v1

Climate Justice Network (n.d.) Hidden statistics: environmental refugees. https://web.archive.org/web/20050223042051/http://www.risingtide.nl/greenpepper/envracism/refugees.html. Accessed 22 March 2020

Coe R (2002) It's the effect size, stupid. a paper presented at the British Educational Research Association annual conference

Connolly K, Chrisafis A, McPherson P, Kirchgaessner S, Haas B, Phillips D, Hunt E, Safi M (2016) Fake News: An Insidious Trend That's Fast Becoming a Global Problem. https://www.theguardian.com/media/2016/dec/02/fake-news-facebook-us-election-around-the-world. Accessed 21 February 2020

Conroy NJ, Rubin VL, Chen Y (2015) Automatic deception detection: Methods for finding fake news. In: Grove A (ed) Proceedings of the Association for Information Science and Technology, vol 52. Association for Information Science and Technology, Silver Spring, Maryland, pp 1 - 4

Croson R, Handy F, Shang J (2009) Keeping Up with The Joneses: The Relationship of Perceived Descriptive Social Norms, Social Information, and Charitable Giving. Nonprofit Management and Leadership 19(4), pp 467 - 489

Dahlinger A, Tiefenbeck V, Ryder B, Gahr B, Fleisch E, Wortmann F (2018) The impact of numerical vs. symbolic eco-driving feedback on fuel consumption – A randomized control field trial. Transportation Research Part D: Transport and Environment 65, pp 375 - 386

Dalén A, Krämer J (2017) Towards a User-Centered Feedback Design for Smart Meter Interfaces to Support Efficient Energy-Use Choices. Business & Information Systems Engineering 59(5), pp 361 - 373

Darby S (2006) The effectiveness of feedback on energy consumption. A Review for DEFRA of the Literature on Metering, Billing and direct Displays

Demarque C, Charalambides L, Hilton DJ, Waroquier L (2015) Nudging Sustainable Consumption: The Use of Descriptive Norms to Promote a Minority Behavior in A

Realistic Online Shopping environment. Journal of Environmental Psychology 43, pp 166 - 174
Denner M-S, Gimpel H, Heger S, Hufnagl C, Schröder J, Thim H, Wöhl M (forthcoming) Das Krankenhaus im Wandel – Herausforderungen und Chancen in der Krankenhauslogistik. In: Gimpel H, Schröder J (eds) Hospital 4.0: schlanke, digital-unterstützte Logistikprozesse in Krankenhäusern, 1st edn. Springer Gabler, Wiesbaden
Denning PJ (1997) A new social contract for research. Communications of the ACM 40(2), pp 132 - 134
Dennis AR, Galletta DF, Webster J (2019) Call for Paper: Special Section on Fake News on the Internet. Journal of Management Information Systems
Deterding S, Dixon D, Khaled R, Nacke L (2011) From Game Design Elements to Gamefulness: Defining Gamification. In: Lugmayr A, Franssila H, Safran C, Hammouda I (eds) Proceedings of the 15th International Academic MindTrek Conference Envisioning Future Media Environments. ACM, New York, NY, pp 9 - 15
Deutsch M, Gerard HB (1955) A Study of Normative and Informational Social Influences Upon Individual Judgment. The Journal of Abnormal and Social Psychology 51(3), pp 629 - 636
Devine PG, Hirt ER, Gehrke EM (1990) Diagnostic and Confirmation Strategies in Trait Hypothesis Testing. Journal of Personality and Social Psychology 58(6), pp 952 - 963
DiFranzo D, Taylor SH, Kazerooni F, Wherry OD, Bazarova NN (2018) Upstanding by Design. In: Mandryk R, Hancock M, Perry M, Cox A (eds) Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI ‘18. ACM Press, New York, New York, USA, pp 1 - 12
Ecker R, Holzer P, Broy V, Butz A (2011) EcoChallenge: a race for efficiency. In: Bylund M, Juhlin O, Fernaeus Y (eds) Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI ‘11. ACM Press, New York, New York, USA, pp 91 - 94
Egan C (1999) How customers interpret and use comparative displays of their home energy use. Proc. European Council for an Energy-Efficient Economy
Elkington J (1994) Towards the Sustainable Corporation: Win-Win-Win Business Strategies for Sustainable Development. California Management Review 36(2), pp 90 - 100
Emmrich V, Döbele M, Bauernhansl T, Paulus-Rohmer D, Schatz A, Weskamp M (2015) Geschäftsmodell-Innovation durch Industrie 4.0: Chancen und Risiken für den

Maschinen- und Anlagenbau. München, Stuttgart: Dr. Wieselhuber & Partner, Fraunhofer IPA

Ericsson E (2001) Independent Driving Pattern Factors and their Influence on Fuel-Use and Exhaust Emission Factors. Transportation Research Part D: Transport and Environment 6(5), pp 325 - 345

Eriksson O, Henderson-Sellers B, Ågerfalk PJ (2013) Ontological and Linguistic Metamodelling Revisited: A Language Use Approach. Information and Software Technology 55(12), pp 2099 - 2124

Erlach K (2010) Wertstromdesign: Der Weg zur schlanken Fabrik, 2nd edn. VDI-Buch. Springer, Berlin, New York

European Parliament (2017) Regulation (EU) 2017/745 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC: OJ L 117. http://data.europa.eu/eli/reg/2017/745/oj. Accessed 27 May 2020

Facebook (2016) Facebook Reports Fourth Quarter and Full Year 2016 Results, Menlo Park, CA

Facebook (2019) Company Info. https://newsroom.fb.com/company-info/. Accessed 3 January 2019

Ferrara E, Varol O, Davis C, Menczer F, Flammini A (2016) The Rise of Social Bots. Communications of the ACM 59(7), pp 96 - 104

Festinger L (1957) A Theory of Cognitive Dissonance. CA: Stanford University Press, Stanford

Fettke P, Loos P (2004) Referenzmodellierungsforschung. WIRTSCHAFTSINFORMATIK 46(5), pp 331 - 340

Fettke P, Loos P (2007) Perspectives on Reference Modeling. In: Fettke P, Loos P (eds) Reference Modeling for Business Systems Analysis. IGI Global, pp 1 - 21

Fettke P, Vom Brocke J (2019) Referenzmodell. In: Gronau N, Becker J, Kliewer N, Leimeister JM, Overhage S (eds) Enzyklopädie der Wirtschaftsinformatik: Online-Lexikon, 11. Auflage. GITO, Berlin

Fischer C (2008) Feedback on Household Electricity Consumption: A Tool for Saving Energy? Energy Efficiency 1(1), pp 79 - 104

Fischer P, Krueger JI, Greitemeyer T, Vogrincic C, Kastenmüller A, Frey D, Heene M, Wicher M, Kainbacher M (2011) The Bystander-Effect: A Meta-Analytic Review on bystander Intervention in Dangerous and Non-Dangerous Emergencies. Psychological Bulletin 137(4), pp 517 - 537

Flintham M, Karner C, Bachour K, Creswick H, Gupta N, Moran S (2018) Falling for Fake News: Investigating the Consumption of News via Social Media. In: Mandryk R, Hancock M, Perry M, Cox A (eds) Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18. ACM Press, New York, New York, USA

Flynn DJ, Nyhan B, Reifler J (2017) The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs About Politics. Political Psychology 38(2), pp 127 - 150

FootPrint Network (2020) Country Overshoot Days 2020. https://www.overshootday.org/newsroom/country-overshoot-days/. Accessed 22 March 2020

Fornell C, Larcker DF (1981) Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. Journal of Marketing Research 18(1), pp 39 - 50

Foster D, Lawson S, Blythe M, Cairns P (2010) Wattsup?: motivating reductions in domestic energy consumption using social networks. In: Hvannberg EÞ, Lárusdóttir MK, Blandford A, Gulliksen J (eds) Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, pp 178 - 187

Frank U (2013) Domain-Specific Modeling Languages: Requirements Analysis and Design Guidelines. In: Reinhartz-Berger I, Sturm A, Clark T, Cohen S, Bettin J (eds) Domain Engineering: Product Lines, Languages, and Conceptual Models. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 133 - 157

Frederick S (2005) Cognitive reflection and decision. The Journal of Economic Perspectives 19(4), pp 25 - 42

Frenkel S, Maheshwari S (2018) Facebook to Let Users Rank Credibility of News. https://www.nytimes.com/2018/01/19/technology/facebook-news-feed.html. Accessed 21 February 2020

Froehlich J, Dillahunt T, Klasnja P, Mankoff J, Consolvo S, Harrison B, Landay JA (2009) UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, Boston, MA, USA, pp 1043 - 1052

Froehlich J, Findlater L, Landay J (2010) The design of eco-feedback technology. In: Mynatt E, Schoner D, Fitzpatrick G, Hudson S, Edwards K, Rodden T (eds) Proceedings of the

28th International Conference on Human Factors in Computing Systems - CHI '10. ACM Press, New York, New York, USA

Garimella K, Morales GDF, Gionis A, Mathioudakis M (2018) Quantifying Controversy on Social Media. ACM Transactions on Social Computing 1(1), pp 1 - 27

Geller ES, Winett RA, Everett PB (1982) Preserving the Environment: New Strategies for Behavior Change. Pergamon Press, New York, NY

Gerber AS, Rogers T (2009) Descriptive Social Norms and Motivation to Vote: Everybody's Voting and so Should You. The Journal of Politics 71(1), pp 178 - 191

Gholami R, Watson R, Hasan H, Molla A, Bjorn-Andersen N (2016) Information Systems Solutions for Environmental Sustainability: How Can We Do More? Journal of the Association for Information Systems 17(8), pp 521 - 536

Gil de Zúñiga H, Jung N, Valenzuela S (2012) Social Media Use for News and Individuals' Social Capital, Civic Engagement and Political Participation. Journal of Computer-Mediated Communication 17(3), pp 319 - 336

Gimpel H, Heger S, Kasper J, Schäfer R (2020a) The Power of Related Articles – Improving Fake News Detection on Social Media Platforms. In: Bui T (ed) Proceedings of the 53rd Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences

Gimpel H, Heger S, Olenberger C, Utz L (2020b) Exploring the use of social norms to improve user's reporting behavior in the fight against fake news. Major revision, 2nd round

Gimpel H, Heger S, Wöhl M (2020c) Sustainable Behavior in Motion: Designing Smart Eco-Driving Feedback Information Systems. Under review

Gimpel H, Hosseini S, Huber R, Probst L, Röglinger M, Faisst U (2018) Structuring Digital Transformation : A Framework of Action Fields and its Application at ZEISS. Journal of Information Technology Theory and Application 19(1), pp 31 - 54

Goldstein BE (2010) Sensation and Perception, 8th edn. Wadsworth Cengage Learning, Belmont, CA

Goldstein NJ, Cialdini RB, Griskevicius V (2008) A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels. Journal of Consumer Research 35(3), pp 472 - 482

Gonzalez-Perez C, Henderson-Sellers B (2008) Metamodelling for Software Engineering. Wiley Publishing

Gottfried J, Shearer E (2016) News Use Across Social Media Platforms 2016. http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/. Accessed 21 February 2020

Graham J, Koo M, Wilson TD (2011) Conserving Energy by Inducing People to Drive Less. Journal of Applied Social Psychology 41(1), pp 106 - 118

Greenhill KM, Oppenheim B (2017) Rumor Has It: The Adoption of Unverified Information in Conflict Zones. International Studies Quarterly 61(3), pp 660 - 676

Gregor (2006) The Nature of Theory in Information Systems. MIS Quarterly 30(3), p 611

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

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

Haakana M, Sillanpää L, Talsi M (eds) (1997) The effect of feedback and focused advice on household energy consumption

Hansen PG, Jespersen AM (2013) Nudge and the Manipulation of Choice. European Journal of Risk Regulation 4(1), pp 3 - 28

Hartmann L, Meudt T, Seifermann S, Metternich J (2018) Value stream method 4.0: holistic method to analyse and design value streams in the digital age. Procedia CIRP 78, pp 249 - 254

Hasher L, Goldstein D, Toppino T (1977) Frequency and The Conference of Referential Validity. Journal of Verbal Learning and Verbal Behavior 16(1), pp 107 - 112

Heger S, Thim H, Gimpel H, Schröder J (2020a) Value Stream Modeling and Notation – A Domain-Specific Modeling Language for the Digital Transformation of Value Streams. Under review

Heger S, Thim H, Wöhl M (forthcoming) Theoretische Grundlagen und Methoden zur Analyse und Gestaltung schlanker, digital-unterstützer Logistikprozesse in Krankenhäusern. In: Gimpel H, Schröder J (eds) Hospital 4.0: schlanke, digital-unterstützte Logistikprozesse in Krankenhäusern, 1st edn. Springer Gabler, Wiesbaden

Heger S, Valett L, Thim H, Schröder J, Gimpel H (2020b) Value Stream Model and Notation – Digitale Transformation von Wertströmen. In: Gronau N, Heine M, Poustcchi K, Krasnova H (eds) WI2020 Zentrale Tracks. GITO Verlag, pp 710 - 724

Henkel C, Kranz J (2018) Pro-Environmental Behavior and Green Information Systems Research – Review, Synthesis and Directions for Future Research. In: Proceedings of the 39th International Conference on Information Systems (ICIS)

Henkel C, Seidler A-R, Kranz J, Fiedler M (2017) How to become a sustainability leader? The role of IS affordances in enabling and triggering sustainability transformations. In: Proceedings of the 38th International Conference on Information Systems (ICIS)

Hermida A (2012) TWEETS AND TRUTH. Journalism Practice 6(5-6), pp 659 - 668

Hevner, March, Park, Ram (2004) Design Science in Information Systems Research. MIS Quarterly 28(1), pp 75 - 105

Hiraoka T, Nishikawa S, Yamabe S, Matsumoto S (2010) Sustainability verification of eco-driving behavior based on driving simulator experiments. In: Proceedings of 17th world congress on intelligent transport systems.

Hiraoka T, Terakado Y, Matsumoto S, Yamabe S (2009) Quantitative evaluation of eco-driving on fuel consumption based on driving simulator experiments. In: Proceedings of the 16th World Congress on Intelligent Transport Systems

Ho DE (2012) Fudging the Nudge: Information Disclosure and Restaurant Grading. Yale Law Journal 122, pp 574 - 688

Horn JL (1965) A rationale and test for the number of factors in factor analysis. Psychometrika 30(2), pp 179 - 185

Housholder EE, LaMarre HL (2014) Facebook Politics: Toward a Process Model for Achieving Political Source Credibility Through Social Media. Journal of Information Technology & Politics 11(4), pp 368 - 382

Huber PJ (1967) The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol 1, pp 221 - 233

Hunt E (2017) ‘Disputed by Multiple Fact-Checkers’: Facebook Rolls Out New Alert to Combat Fake News. https://www.theguardian.com/technology/2017/mar/22/facebook-fact-checking-tool-fake-news. Accessed 22 January 2019

Huotari K, Hamari J (2012) Defining Gamification - A Service Marketing Perspective. In: Lugmayr A (ed) Proceedings of the 16th International Academic MindTrek Conference. ACM, New York, NY, pp 17 - 22

Ilgen DR, Fisher CD, Taylor MS (1979) Consequences of individual feedback on behavior in organizations. Journal of Applied Psychology 64(4), pp 349 - 371

IPCC (2014) Climate change: Impacts, adaptation, and vulnerability IPCC working group II contribution to AR5 report. https://www.ipcc.ch/report/ar5/wg2/. Accessed 22 March 2020

Jacobson RP, Mortensen CR, Cialdini RB (2011) Bodies Obliged and Unbound: Differentiated Response Tendencies for Injunctive and Descriptive Social Norms. Journal of Personality and Social Psychology 100(3), pp 433 - 448

Jain RK, Taylor JE, Peschiera G (2012) Assessing eco-feedback interface usage and design to drive energy efficiency in buildings. Energy and Buildings 48, pp 8 - 17

Jamieson A, Solon O (2016) Facebook to Begin Flagging Fake News in Response to Mounting Criticism. https://www.theguardian.com/technology/2016/dec/15/facebook-flag-fake-news-fact-check. Accessed 21 February 2020

Jamson SL, Hibberd DL, Jamson AH (2015) Drivers' ability to learn eco-driving skills; effects on fuel efficient and safe driving behaviour. Transportation Research Part C: Emerging Technologies 58, pp 657 - 668

Kamilaris A, Pitsillides A (2016) Mobile Phone Computing and the Internet of Things: A Survey. IEEE Internet of Things Journal 3(6), pp 885 - 898

Kaplan AM, Haenlein M (2010) Users of The World, Unite!: The Challenges and Opportunities of Social Media. Business Horizons 53(1), pp 59 - 68

Karlin B, Zinger JF, Ford R (2015) The effects of feedback on energy conservation: A meta-analysis. Psychological Bulletin 141(6), pp 1205 - 1227

Kelkar O, Heger R, Dao D-K Studie Industrie 4.0–Eine Standortbestimmung der Automobil- und Fertigungsindustrie. http://www.mhp.com/fileadmin/mhp.de/assets/studien/MHP-Studie_Industrie4.0_V1.0.pdf. Accessed 4 April 2019

Kim A, Dennis AR (2019) Says Who? The Effects of Presentation Format and Source Rating on Fake News in Social Media. MIS Quarterly 43(3), 1025-1039

Kim A, Moravec PL, Dennis AR (2019) Combating Fake News on Social Media with Source Ratings: The Effects of User and Expert Reputation Ratings. Journal of Management Information Systems 36(3), pp 931 - 968

Klevers T (2009) Wertstrom-Mapping und Wertstrom-Design: Verschwendung erkennen - Wertschöpfung steigern, 1st edn. mi-Wirtschaftsbuch, München

Kluger A, DeNisi A (1996) The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin 119(2), pp 254 - 284

Koch S (2015) Einführung in das Management von Geschäftsprozessen: Six Sigma, Kaizen und TQM, 2nd edn. Springer Vieweg, Berlin

König UM, Linhart A, Röglinger M (2019) Why do business processes deviate? Results from a Delphi study. Business Research 12(2), pp 425 - 453

Koriat A, Lichtenstein S, Fischhoff B (1980) Reasons for Confidence. Journal of Experimental Psychology: Human Learning & Memory 6(2), pp 107 - 118

Kormos C, Gifford R, Brown E (2014) The Influence of Descriptive Social Norm Information on Sustainable Transportation Behavior. Environment and Behavior 47(5), pp 479 - 501

Kranz J, Kolbe LM, Koo C, Boudreau M-C (2015) Smart energy: where do we stand and where should we go? Electronic Markets 25(1), pp 7 - 16

Kriegel J, Jehle F, Moser H, Tuttle-Weidinger L (2016) Patient logistics management of patient flows in hospitals: A comparison of Bavarian and Austrian hospitals. International Journal of Healthcare Management 9(4), pp 257 - 268

Kroese FM, Marchiori DR, Ridder DTD de (2016) Nudging healthy food choices: a field experiment at the train station. Journal of public health 38(2), 133-137

Kuem J, Ray S, Siponen M, Kim SS (2017) What Leads to Prosocial Behaviors on Social Networking Services: A Tripartite Model. Journal of Management Information Systems 34(1), pp 40 - 70

Kuhlang P, Edtmayr T, Sihn W (2011) Methodical approach to increase productivity and reduce lead time in assembly and production-logistic processes. CIRP Journal of Manufacturing Science and Technology 4(1), pp 24 - 32

Kurani KS, Stillwater T, Jones M, Caperello N (2013) Ecodrive I-80: A Large Sample Fuel Economy Feedback Field Test Final Report: Report ITS-RR-13-15, Davis, CA

LaBrie JW, Hummer JF, Neighbors C, Larimer ME (2010) Whose Opinion Matters? The Relationship between Injunctive Norms and Alcohol Consequences in College Students. Addictive Behaviors 35(4), pp 343 - 349

Lange F, Dewitte S (2019) Measuring pro-environmental behavior: Review and recommendations. Journal of Environmental Psychology 63, pp 92 - 100

Lárusdóttir EB, Ulfarsson GF (2014) Effect of Driving Behavior and Vehicle Characteristics on Energy Consumption of Road Vehicles Running on Alternative Energy Sources. International Journal of Sustainable Transportation 9(8), pp 592 - 601

Latan B, Nida S (1981) Ten Years of Research on Group Size and Helping. Psychological Bulletin 89(2), pp 308 - 324

Latané B, Darley JM (1970) The Unresponsive Bystander: Why Doesn't He Help? Century Psychology Series. Appleton-Century Crofts, New York

Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, Metzger MJ, Nyhan B, Pennycook G, Rothschild D, Schudson M, Sloman SA, Sunstein CR, Thorson EA, Watts DJ, Zittrain JL (2018) The Science of Fake News. Science 359(6380), pp 1094 - 1096

Lee AS, Thomas M, Baskerville RL (2015) Going back to basics in design science: from the information technology artifact to the information systems artifact. Information Systems Journal 25(1), pp 5 - 21

Lee CF, Lee JC, Lee AC (2000) Statistics for Business and Financial Economics, 2. ed. World Scientific, Singapore

Legner C, Eymann T, Hess T, Matt C, Böhmann T, Drews P, Mädche A, Urbach N, Ahlemann F (2017) Digitalization: Opportunity and Challenge for the Business and Information Systems Engineering Community. Business & Information Systems Engineering 59(4), pp 301 - 308

Legner C, Pentek T, Otto B (2020) Accumulating Design Knowledge with Reference Models: Insights from 12 Years' Research into Data Management. Journal of the Association for Information Systems 21(3), pp 735 - 770

Leonard TC (2008) Richard H. Thaler, Cass R. Sunstein, Nudge: Improving decisions about health, wealth, and happiness. Constitutional Political Economy 19(4), pp 356 - 360

Lewin K (1936) Principles of Topological Psychology. Mcgraw-Hill Book Company, New York, NY

Lodderstedt T, Basin D, Doser J (eds) (2002) SecureUML: A UML-Based Modeling Language for Model-Driven Security. ≪UML≫ 2002 — The Unified Modeling Language. Springer Berlin Heidelberg

Loeser F (2013) Green IT and Green IS: Definition of constructs and overview of current practices. In: Proceedings of the 19th Americas Conference on Information Systems (AMCIS)

Lord CG, Lepper MR, Preston E (1984) Considering the opposite: A corrective strategy for social judgment. Journal of Personality and Social Psychology 47(6), pp 1231 - 1243

Maasberg M, Ayaburi E, Liu C, Au Y (2018) Exploring the Propagation of Fake Cyber News: An Experimental Approach. In: Bui T (ed) Proceedings of the 51st Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences

Maedche A, Morana S, Schacht S, Werth D, Krumeich J (2016) Advanced User Assistance Systems. Business & Information Systems Engineering 58(5), pp 367 - 370

Magana VC, Organero MM (2013) GAFU: A game to save fuel using social networks. In: 2013 International Conference on Connected Vehicles and Expo (ICCVE). IEEE, pp 151 - 157

Maksl A, Ashley S, Craft S (2015) Measuring news media literacy. Journal of Media Literacy Education 6(3), pp 29 - 45

March ST, Smith GF (1995) Design and natural science research on information technology. Decision Support Systems 15(4), pp 251 - 266

Martin EW, Boriboonsomsin K, Chan ND, Williams N, Shaheen SA, Barth M (2013) Dynamic ecodriving in northern california: A study of survey and vehicle operations data from an ecodriving feedback device. In: 92nd Annual Meeting of the Transportation Research Board

Martínez-Jurado PJ, Moyano-Fuentes J (2014) Lean Management, Supply Chain Management and Sustainability: A Literature Review. Journal of Cleaner Production 85, pp 134 - 150

McAfee A (2019) More from less: The surprising story of how we learned to prosper using fewer resources - and what happens next. Simon & Schuster, London

Melville NP (2010) Information Systems Innovation for Environmental Sustainability. MIS Quarterly 34(1), pp 1 - 21

Messerli P, Murniningtyas E, Eloundou-Enyegue P, Foli EG, Furman E, Glassman A, Hernández Licona G, Kim EM, Lutz W, Moatti J-P (2019) Global Sustainable Development Report: The Future is Now - Science for Achieving Sustainable Development. http://pure.iiasa.ac.at/id/eprint/16067/1/24797GSDR_report_2019.pdf. Accessed 27 March 2020

Metternich J, Müller M, Meudt T, Schaede C (2017) Lean 4.0 – zwischen Widerspruch und Vision. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 112(5), pp 346 - 348

Metzler C, Dollinger K, Kurzeder S, Israel B (2013) A method for displaying a gear change recommendation(DE 102013217194 A1). http://www.google.com/patents/DE102013217194A1?cl=en

Meudt T, Metternich J, Abele E (2017) Value stream mapping 4.0: Holistic examination of value stream and information logistics in production. CIRP Annals 66(1), pp 413 - 416

Meyer UB, Creux SEM, Weber Marin Silva AK (2005) Grafische Methoden der Prozessanalyse: Für Design und Optimierung von Produktionssystemen. Hanser, München

Michon JA (1985) A Critical View of Driver Behavior Models: What Do We Know, What Should We Do. In: Evans L, Schwing RC (eds) Human Behavior and Traffic Safety. Springer US, Boston, MA, pp 485 - 520

Mills J (2019) Improving The 1957 Version of Dissonance Theory. In: Harmon-Jones E (ed) Cognitive dissonance: Reexamining a pivotal theory in psychology (2nd ed.). American Psychological Association, Washington, pp 27 - 39

Moravec PL, Minas R, Dennis AR (2019) Fake News on Social Media: People Believe What They Want to Believe When It Makes No Sense At All. MIS Quarterly 43(4), pp 1343 - 1360

Mosseri A (2016) News Feed FYI: Addressing Hoaxes and Fake News. https://newsroom.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/. Accessed 22 January 2019

Murray A (2016) How to Report Fake News to Social Media. https://www.bbc.com/news/38053324. Accessed 4 January 2019

Mustafaraj E, Metaxas PT (2017) The Fake News Spreading Plague: Was it Preventable? In: Boldi P, Kinder-Kurlanda K, Fox P, McGuinness D, Poirer L (eds) Proceedings of the 2017 ACM Web Science Conference. ACM Association for Computing Machinery, New York, NY, USA, pp 235 - 240

Nachtmann H, Pohl EA (2009) The State of Healthcare Logistics: Cost and Quality Improvement Opportunities, University of Arkansas

Neighbors C, Lostutter TW, Whiteside U, Fossos N, Walker DD, Larimer ME (2007) Injunctive Norms and Problem Gambling Among College Students. Journal of Gambling Studies 23(3), pp 259 - 273

Newman N, Fletcher R, Kalogeropoulos A, Levy DAL, Nielsen RK (2017) Reuters Institute Digital News Report 2017

Newman N, Fletcher R, Kalogeropoulos A, Levy DAL, Nielsen RK (2018) Digital News Report 2018. media.digitalnewsreport.org/wp-content/uploads/2018/06/digital-news-report-2018.pdf?x89475. Accessed 1 November 2018

Nickerson RS (1998) Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. Review of General Psychology 2(2), pp 175 - 220

Nigbur D, Lyons E, Uzzell D (2010) Attitudes, Norms, Identity and Environmental Behaviour: Using an Expanded Theory of Planned Behaviour to Predict Participation in A Kerbside Recycling programme. The British Journal of Social Psychology 49(2), pp 259 - 284

Nordstrom G, Sztipanovits J, Karsai G, Ledeczi A (1999) Metamodeling - Rapid Design and Evolution of Domain-Specific Modeling Environments. In: Proceedings of the 1999 IEEE Conference on Engineering of Computer-Based Systems. IEEE Computer Society, Nashville, Tennessee, pp 68 - 74

Nyhan B, Reifler J (2010) When Corrections Fail: The Persistence of Political Misperceptions. Political Behavior 32(2), pp 303 - 330

Object Management Group (2011) Business Process Model and Notation (BPMN): Version 2.0. http://www.omg.org/spec/BPMN/2.0. Accessed 25 July 2019

Object Management Group (2015) OMG Unified Modeling Language (OMG UML): Version 2.5. https://www.omg.org/spec/UML/2.5/PDF. Accessed 24 July 2019

Oltermann P (2018) Tough New German Law Puts Tech Firms and Free Speech in Spotlight. https://www.theguardian.com/world/2018/jan/05/tough-new-german-law-puts-tech-firms-and-free-speech-in-spotlight. Accessed 21 February 2020

Oxford Internet Institute, University of Oxford (2016) Resource For Understanding Political Bots. https://www.oii.ox.ac.uk/blog/resource-for-understanding-political-bots/. Accessed 21 February 2020

Paay J, Kjeldskov J, Skov M, Pathmanathan R, Pearce J (2013) Promoting pro-environmental behaviour. In: Shen H, Smith R, Paay J, Calder P, Wyeld T (eds) OzCHI 2013: Proceedings of the 25th Australian Computer-Human Interaction Conference. ACM, New York, pp 235 - 244

Park HS, Smith SW (2007) Distinctiveness and Influence of Subjective Norms, Personal Descriptive and Injunctive Norms, and Societal Descriptive and Injunctive Norms on Behavioral Intent: A Case of Two Behaviors Critical to Organ Donation. Human Communication Research 33(2), pp 194 - 218

Parmiggiani E, Monteiro E (2018) Shifting Baselines? Recommendations for Green IS. In: Proceedings of the 39th International Conference on Information Systems (ICIS)

Parrique T, Barth J, Briens F, Kerschner C, Kraus-Pol A, Kuokkanen A, Spangenberg JH (2019) Decoupling debunked: Evidence and arguments against green growth as a sole strategy for sustainability. eeb.org/decoupling-debunked. Accessed 8 June 2020

Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S (2007) A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24(3), pp 45 - 77

Pempek TA, Yermolayeva YA, Calvert SL (2009) College Students' Social Networking Experiences on Facebook. Journal of Applied Developmental Psychology 30(3), pp 227 - 238

Pennycook G, Cannon TD, Rand DG (2018) Prior Exposure Increases Perceived Accuracy of Fake News. Journal of Experimental Psychology: General 147(12), pp 1865 - 1880

Pentina I, Tarafdar M (2014) From "information" to "knowing": Exploring the role of social media in contemporary news consumption. Computers in Human Behavior 35, pp 211 - 223

Peter L, Diane L, Marc A, David F, Betty F, McGrath JE (1972) Bystander Effect in A Demand-without-Threat Situation. Journal of Personality and Social Psychology 24(2), pp 166 - 171

Petri CA (1962) Kommunikation mit Automaten

Petty RE, Cacioppo JT (1986) Elaboration Likelihood Model. In: Berkowitz L (ed) Advances in Experimental Social Psychology. Academic Press, San Diego, USA, pp. 123–205

Pinsonneault A, Kraemer KL (2002) Exploring the Role of Information Technology in Organizational Downsizing: A Tale of Two American Cities. Organization Science 13(2), pp 191 - 208

Pries-Heje J, Baskerville R, Venable J (2008) Strategies for Design Science Research Evaluation. In: ECIS 2008 Proceedings

Rafael M, Sanchez M, Mucino V, Cervantes J, Lozano A (2006) Impact of Driving Styles on Exhaust Emissions and Fuel Economy From a Heavy-Duty Truck: Laboratory tests. International Journal of Heavy Vehicle Systems 13(1/2), p 56

Reno RR, Cialdini RB, Kallgren CA (1993) The Transsituational Influence of Social Norms. Journal of Personality and Social Psychology 64(1), pp 104 - 112

Revelle W, Revelle MW (n.d.) Package 'psych'

Rice B (1982) The Hawthorne Defect: Persistence of a Flawed Theory. Psychology Today 16(2), pp 70 - 74

Riche Y, Dodge J, Metoyer RA (2010) Studying always-on electricity feedback in the home. In: Mynatt E, Schoner D, Fitzpatrick G, Hudson S, Edwards K, Rodden T (eds) Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10. ACM Press, New York, New York, USA

Rimal RN, Real K (2003) Understanding the Influence of Perceived Norms on Behaviors. Communication Theory 13(2), pp 184 - 203

Robinson E, Fleming A, Higgs S (2014) Prompting Healthier Eating: Testing The Use of Health and Social Norm Based Messages. Health psychology : Official Journal of the Division of Health Psychology, American Psychological Association 33(9), pp 1057 - 1064

Roh P, Kunz A, Wegener K (2019) Information stream mapping: Mapping, analysing and improving the efficiency of information streams in manufacturing value streams. CIRP Journal of Manufacturing Science and Technology 25, pp 1 - 13

Roser C, Lorentzen K, Deuse J (2014) Reliable Shop Floor Bottleneck Detection for Flow Lines through Process and Inventory Observations. Procedia CIRP 19, pp 63 - 68

Ross AD, Jayaraman V (2009) Strategic Purchases of Bundled Products in a Health Care Supply Chain Environment. Decision Sciences 40(2), pp 269 - 293

Rubin VL, Chen Y, Conroy NJ (2015) Deception detection for news: Three types of fakes. In: Grove A (ed) Proceedings of the Association for Information Science and Technology, vol 52. Association for Information Science and Technology, Silver Spring, Maryland, pp 1 - 4

Saboohi Y, Farzaneh H (2009) Model for developing an eco-driving strategy of a passenger vehicle based on the least fuel consumption. Applied Energy 86(10), pp 1925 - 1932

Sarker S, Chatterjee S, Xiao X, Elbanna A (2019) The Sociotechnical Axis of Cohesion for the IS Discipline: Its Historical Legacy and its Continued Relevance. MIS Quarterly 43(3), pp 695 - 719

Sarkis J, Koo C, Watson RT (2013) Green information systems & technologies – this generation and beyond: Introduction to the special issue. Information Systems Frontiers 15(5), pp 695 - 704

Schermann M, Böhmann T, Krcmar H (2009) Explicating Design Theories with Conceptual Models: Towards a Theoretical Role of Reference Models. In: Becker J, Krcmar H, Niehaves B (eds) Wissenschaftstheorie und gestaltungsorientierte Wirtschaftsinformatik. Physica-Verlag HD, Heidelberg, pp 175 - 194

Schmehl K (2017) 7 der 10 erfolgreichsten Artikel über Angela Merkel auf Facebook sind Fake News. https://www.buzzfeed.com/karstenschmehl/die-top-fake-news-ueber-angela-merkel. Accessed 22 January 2019

Schultz PW (2014) Strategies for Promoting Proenvironmental Behavior. European Psychologist 19(2), pp 107 - 117

Schultz PW, Khazian AM, Zaleski AC (2008) Using Normative Social Influence to Promote Conservation Among Hotel Guests. Social Influence 3(1), pp 4 - 23

Schultz PW, Nolan JM, Cialdini RB, Goldstein NJ, Griskevicius V (2007) The Constructive, Destructive, and Reconstructive Power of Social Norms. Psychological Science 18(5), pp 429 - 434

Schwarz N, Sanna LJ, Skurnik I, Yoon C (2007) Metacognitive Experiences and the Intricacies of Setting People Straight: Implications for Debiasing and Public Information Campaigns. In: Zanna MP (ed) Advances in Experimental Social Psychology, 1st edn., vol 39. Academic Press, pp 127 - 161

Sebastian I, Ross J, Beath C, Mocker M, Moloney K, Fonstad N (2017) How big old companies navigate digital transformation. MIS Quarterly Executive 16(3), pp 197 - 213

Seidel S, Bharati P, Fridgen G, Watson RT, Albizri A, Boudreau M-C, Butler T, Chandra Kruse L, Guzman I, Karsten H, Lee H, Melville N, Rush D, Toland J, Watts S (2017) The sustainability imperative in information systems research. Communications of the Association for Information Systems 40, pp 40 - 52

Shahab Q, Terken J, Eggen B (2013) Development of a questionnaire for identifying driver’s personal values in driving. In: Terken J (ed) Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI ‘13. ACM Press, New York, New York, USA, pp 202 - 208

Shaheen SA, Cohen AP (2013) Carsharing and Personal Vehicle Services: Worldwide Market Developments and Emerging Trends. International Journal of Sustainable Transportation 7(1), pp 5 - 34

Shao C, Ciampaglia GL, Flammini A, Menczer F (2016) Hoaxy: A Platform for Tracking Online Misinformation. In: Bourdeau J, Hendler JA, Nkambou RN, Horrocks I, Zhao BY (eds) WWW’16 companion: Proceedings of the 25th International Conference on World

Wide Web : May 11-15, 2016, Montreal, Canada. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, pp 745 - 750

Shu K, Sliva A, Wang S, Tang J, Liu H (2017) Fake News Detection on Social Media: A Data Mining Perspective. ACM SIGKDD Explorations Newsletter 19(1), pp 22 - 36

Shukla S, Lyons T (2017) Blocking Ads from Pages that Repeatedly Share False News. https://newsroom.fb.com/news/2017/08/blocking-ads-from-pages-that-repeatedly-share-false-news/. Accessed 22 January 2019

Siero FW, Bakker AB, Dekker GB, Van Den Burg MTC (1996) Changing organizational energy consumption behaviour through comparative feedback. Journal of Environmental Psychology 16(3), pp 235 - 246

Siero S, Boon M, Kok G, Siero F (1989) Modification of Driving Behavior in a Large Transport Organization: A field experiment. Journal of Applied Psychology 74(3), pp 417 - 423

Silverman C (2016) Here Are 50 Of The Biggest Fake News Hits On Facebook From 2016. https://www.buzzfeed.com/craigsilverman/top-fake-news-of-2016. Accessed 22 January 2019

Simon HA (1979) Rational Decision Making in Business Organizations. The American Economic Review 69(4), pp 493 - 513

Simon HA (1996) The sciences of the artificial. MIT press

Sivak M, Schoettle B (2012) Eco-driving: Strategic, tactical, and operational decisions of the driver that influence vehicle fuel economy. Transport Policy 22, pp 96 - 99

Sokolova M, Lapalme G (2009) A systematic analysis of performance measures for classification tasks. Information Processing & Management 45(4), pp 427 - 437

Sonnenberg C, Vom Brocke J (2012) Evaluation Patterns for Design Science Research Artefacts. In: Helfert M, Donnellan B (eds) Practical Aspects of Design Science, vol 286. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 71 - 83

Staats H, Harland P, Wilke H am (2004) Effecting durable change: A team approach to improve environmental behavior in the household. Environment and Behavior 36(3), pp 341 - 367

Statista (2018a) Most popular Facebook activities in the United States as of October 2017. https://www.statista.com/statistics/275788/share-of-facebook-user-activities/. Accessed 9 June 2019

Statista (2018b) Total number of Facebook users in Germany in January 2017, by age group and gender (in millions). https://www.statista.com/statistics/867622/number-of-facebook-users-in-germany-by-age-and-gender/. Accessed 21 September 2018

Sunstein CR (2014) Nudging: A Very Short Guide. Journal of Consumer Policy 37(4), pp 583 - 588

Thaler RH, Sunstein CR (2003) Libertarian Paternalism. American Economic Review 93(2), pp 175 - 179

Thaler RH, Sunstein CR (2008) Nudge: improving decisions about health, wealth, and happiness. Yale University Press,, New Haven

The Editorial Board (2016) Facebook and the Digital Virus Called Fake News. https://www.nytimes.com/2016/11/20/opinion/sunday/facebook-and-the-digital-virus-called-fake-news.html. Accessed 21 February 2020

Thim H, Heger S (forthcoming) Referenzmodell „Materiallogistik und Medikalprodukteversorgung“. In: Gimpel H, Schröder J (eds) Hospital 4.0: schlanke, digital-unterstützte Logistikprozesse in Krankenhäusern, 1st edn. Springer Gabler, Wiesbaden

Tomanek DP, Schröder J (2018) Value Added Heat Map: Eine Methode zur Visualisierung von Wertschöpfung. Springer Gabler, Wiesbaden

Tulusan J, Staake T, Fleisch E (2012a) Direct or indirect sensor enabled eco-driving feedback: Which preference do corporate car drivers have? In: Wang J (ed) 2012 3rd IEEE International Conference on the Internet of Things (IOT 2012). IEEE, Piscataway, NJ, pp 39 - 46

Tulusan J, Staake T, Fleisch E (2012b) Providing Eco-Driving Feedback to Corporate Car Drivers. In: Proceedings of the 2012 ACM Conference on Ubiquitous Computing. ACM, New York, NY, pp 212 - 215

Twitter International Company (2018) Q3 2018 Letter to Shareholders. https://s22.q4cdn.com/826641620/files/doc_financials/2018/q3/Q3-2018-Shareholder-Letter.pdf. Accessed 3 January 2019

Uckelmann D (2014) Wertstromorientierte Informationsflüsse für Industrie 4.0—Kernprozesse und Gestaltungsvariablen. Industrie Management 6(2014), pp 13 - 17

Ueno T, Sano F, Saeki O, Tsuji K (2006) Effectiveness of an energy-consumption information system on energy savings in residential houses based on monitored data. Applied Energy 83(2), pp 166 - 183

United Nations (2015) Transforming our world: the 2030 Agenda for Sustainable Development. https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E. Accessed 7 April 2020

Vagg C, Brace CJ, Hari D, Akehurst S, Poxon J, Ash L (2013) Development and Field Trial of a Driver Assistance System to Encourage Eco-Driving in Light Commercial Vehicle Fleets. IEEE Transactions on Intelligent Transportation Systems 14(2), pp 796 - 805

van der Aalst WMP (2013) Business Process Management: A Comprehensive Survey. ISRN Software Engineering 2013(1), pp 1 - 37

van der Linden S, Leiserowitz A, Rosenthal S, Maibach E (2017) Inoculating the Public against Misinformation about Climate Change. Global Challenges 1(2)

van der Molen HH, Botticher AMT (1987) Risk models for traffic participants: A concerted effort for theoretical operationalizations. In: Rothengatter JA, Bruin RAd (eds) Road Users and Traffic Safety. Van Gorcum, Assen/Maastricht, pp 61 - 82

van der Voort M, Dougherty MS, van Maarseveen M (2001) A Prototype Fuel-Efficiency Support Tool. Transportation Research Part C: Emerging Technologies 9(4), pp 279 - 296

VDI Association of German Engineers e. V. (VDI) (2012) Ganzheitliche Produktionssysteme: Grundlagen, Einführung und Bewertung

VDI Association of German Engineers e. V. (VDI) (2020) Standards setzen - auf dem aktuellen Stand der Technik. https://www.vdi.de/richtlinien. Accessed 10.05.20

Venable J, Pries-Heje J, Baskerville R (2012) A Comprehensive Framework for Evaluation in Design Science Research. In: Peffers K, Rothenberger M, Kuechler B (eds) Design Science Research in Information Systems. Advances in Theory and Practice. Springer-Verlag Berlin Heidelberg, Heidelberg, pp 423 - 438

Venable J, Pries-Heje J, Baskerville R (2016) FEDS: a Framework for Evaluation in Design Science Research. European Journal of Information Systems 25(1), pp 77 - 89

Venkatesh V, Brown SA, Bala H (2013) Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems. MIS Quarterly 37(1), pp 21 - 54

Vial G (2019) Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems 28(2), pp 118 - 144

Vicario MD, Quattrociocchi W, Scala A, Zollo F (2019) Polarization and Fake News: Early Warning of Potential Misinformation Targets. ACM Transactions on the Web 13(2), pp 1 - 22

Voelpel SC, Eckhoff RA, Förster J (2008) David against Goliath?: Group Size and Bystander Effects in Virtual Knowledge Sharing. Human Relations 61(2), pp 271 - 295

Volland J, Fügener A, Schoenfelder J, Brunner JO (2017) Material logistics in hospitals: A literature review. Omega 69, pp 82 - 101

Vom Brocke J, Watson RT, Dwyer C, Elliot S, Melville N (2013) Green Information Systems: Directives for the IS Discipline. Communications of the Association for Information Systems 33

Vosoughi S, Roy D, Aral S (2018) The Spread of True and False News Online. Science 359(6380), pp 1146 - 1151

Wallace L, Keil M, Rai A (2004) How Software Project Risk Affects Project Performance: An Investigation of the Dimensions of Risk and an Exploratory Model. Decision Sciences 35(2), pp 289 - 321

Walls JG, Widmeyer GR, El Sawy OA (1992) Building an Information System Design Theory for Vigilant EIS. Information Systems Research 3(1), pp 36 - 59

Walsham G (2017) ICT4D research: reflections on history and future agenda. Information Technology for Development 23(1), pp 18 - 41

Wand Y, Weber R (2002) Research Commentary: Information Systems and Conceptual Modeling—A Research Agenda. Information Systems Research 13(4), pp 363 - 376

Warzel C, Mac R (2018) Twitter Just Launched A Midterm Elections Page And It's Already Full Of Garbage. https://www.buzzfeednews.com/article/charliewarzel/twitter-just-launched-a-midterms-page-and-its-already. Accessed 3 January 2019

Wathen CN, Burkell J (2002) Believe It or Not: Factors Influencing Credibility on The Web. Journal of the American Society for Information Science and Technology 53(2), pp 134 - 144

Watson RT, Boudreau M-C, Chen AJ (2010) Information systems and environmentally sustainable development: energy informatics and new directions for the IS community. MIS Quarterly 34(1), pp 23 - 38

Weedon J, Nuland W, Stamos A (2017) Information Operations and Facebook. https://fbnewsroomus.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf. Accessed 22 January 2019

Weeks BE, Friedenberg LM, Southwell BG, Slater JS (2012) Behavioral consequences of conflict-oriented health news coverage: the 2009 mammography guideline controversy and online information seeking. Health Communication 27(2), pp 158 - 166

Weinmann M, Schneider C, Vom Brocke J (2016) Digital Nudging. Business & Information Systems Engineering 58(6), pp 433 - 436

Wenzel M (2004) The Social Side of Sanctions: Personal and Social Norms as Moderators of Deterrence. Law and Human Behavior 28(5), pp 547 - 567

White H (1980) A Heteroskedasticity-Consistent Covariance Matrix Estimator and A Direct Test for Heteroskedasticity. Econometrica: Journal of the Econometric Society, pp 817 - 838

Wiley J (2005) A fair and balanced look at the news: What affects memory for controversial arguments? Journal of Memory and Language 53(1), pp 95 - 109

Womack JP, Jones DT (2013) Lean Thinking: Banish Waste and Create Wealth in Your Corporation. Simon & Schuster, Limited, London

Wong RYM, Cheung CMK, Xiao B (2016) Combating Online Abuse: What Drives People to Use Online Reporting Functions on Social Networking Sites. In: 2016 49th Hawaii International Conference on System Sciences (HICSS), pp 415 - 424

World Commission on Environment and Development (WCED) (1987) Our Common Future. https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf. Accessed 9 June 2020

Yang W, Leon PG, Scott K, Chen X, Acquisti A, Cranor LF (2013) Privacy nudges for social media: an exploratory Facebook study. In: Schwabe D (ed) Proceedings of the 22nd International Conference on World Wide Web Companion. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, pp 763 - 770

Appendix A Designing Eco-Driving Feedback Information Systems – Results from a Field Study to Improve Driving Behavior

Throughout the iterative search process, an earlier version of the design artifact has been launched within the field study. The release did not include the aspects regarding normative, i.e., community feedback (DP 3-4). Technically the versions are identical, except that the mature version additionally provides norm activation. Initially, we did not expect normative feedback to be effective, as the effectiveness of social comparisons is questionable from a psychological point of view. On the one hand, Froehlich et al. (2010) outline that depending on the individual value system, SN do or do not influence pro-environmental behavior (cf. rational-economic model vs. norm-activation model). Some studies found evidence that social feedback is more effective than feedback merely on the own performance (Paay et al. 2013; Siero et al. 1996; Goldstein et al. 2008), other studies could not find significant improvements through social feedback (Haakana et al. 1997; Egan 1999; Schultz et al. 2007).

Similarly, the results by Bergquist (2020) indicate that with regards to pro-environmental behavior, a better-than-average effect is observable, meaning that people are overestimating their efforts with regards to environmental behavior in comparison to their peers. Further, insights from different studies suggest that social comparison can result in the so-called *boomerang effect*. The effect describes the phenomena that people improve their behavior as long as other people are doing better but are likely to impair their behavior when most people are doing worse (Schultz et al. 2007; Allcott 2011; Fischer 2008). Also, Schultz (2014) argues that depending on the potential benefits of one's behavior, among others, education is enough to foster pro-environmental behavior if the benefits are high (e.g., cost savings from lower fuel consumption). But if a person perceives the benefits as low, education is not enough. Instead, SN foster pro-environmental behavior. From the authors' point of view, a statement about the perceived benefits of a person is ex-ante impossible, and also the individual's value system is not observable. Therefore the design process consisted of at least two iterations: the first iteration excluding normative feedback (DP 3-4) and the second iteration, including normative feedback. Results indicate that the earlier version (1st iteration) seemed not to be sufficient to motivate more fuel-efficient driving behavior on average.

In total, 42 participants (different from those using the mature version) used the app in the earlier version to decrease their consumption. We also calculate the difference of the mean consumption for the first two trips and the average consumption of the six remaining trips. They achieved fuel savings about 0.08 l/100km on average (median -0.03 l/100km, i.e., an increase in fuel consumption). The differences are then also tested with the statistical H_0: Fuel-saving less or equal to zero. The one-sample t-test does not allow to reject H_0 (p-value = 0.285). There has been no significant fuel-saving effect. Nevertheless, we also perform a Wilcoxon Signed-Rank test, which led to comparable results. We could not reject the null hypotheses (p-value = 0.705). For completeness, we also evaluated the size of the effect. Thus, we calculated the effect size. Concerning the earlier version, Cohen's d is 0.04, which is almost no effect. That altogether differs from the fuel savings of the participants using the mature version (Table 20).

	n	***Mean savings***	***Median savings***	***p-value (t-test)***	***p-value (Wilcoxon Signed-Rank test)***	***Cohen's d***
Early version (1st group)	42	0.08 l/100km	-0.03 l/100km	0.285	0.705	0.04
Mature version (2nd group)	40	0.35 l/100km	0.14 l/100km	0.022	0.025	0.21

Table 20: Significance and size of the fuel-saving effect for both versions

Note: H_0 = Fuel saving is less or equal to zero

The participants' characteristics of both groups are comparable as a Chi-squared test of homogeneity is not significant for all reported characteristics (demographics, car characteristics, environmental attitudes). The difference in means between both groups is 0.27 l/100km. We apply a t-test to investigate the differences between both groups. Based on our previous results, we assume the savings in the second group are higher than in the first group. Thus, the statistical H_0 is that the fuel savings in the first group are higher or equal to the savings in the second group. We cannot reject H_0 as the p-value of the t-test is 0.104. As none of the two samples (savings in the first group and savings in the second group) are normally distributed, we also apply a Wilcoxon-Mann-Whitney test. The result leads to the conclusion that, indeed, the savings in the second group are significantly higher than in the first group (p-value = 0.044). In line with these findings, the Kolmogorov-Smirnov test also indicates that both samples are drawn from different distributions, and the cumulative distribution function in the first group lies significantly above the cumulative distribution function in the second group (p-value = 0.036). The result implies that the mature version, which has been applied to the second group, is more effective than the earlier version. Furthermore, Cohen's d between both is 0.28, and thus we observe a small effect (Table 21).

	Difference of mean fuel savings	***p-value of t-test***	***p-value of Wilcoxon Signed-Rank test***	***p-value of Kolmogorov-Smirnov test***	***Cohen's d***
1st vs. 2nd	0.27 l/100km	0.104	0.044	0.036	0.28

Table 21: Comparison of the effectiveness of both versions

Note: H_0 = Fuel savings in group E are greater or equal to savings in group S

We also performed the analysis mentioned above on all available trips per driver to investigate the fuel-saving effect over the full time each participant used the app. Results remain overall the same. The effect in the 1st group is neither substantial nor significant.

Appendix B Improving Users' Fake News Recognition in Social Media – an Online Experiment

Appendix B.1 Procedure of the Online Experiment

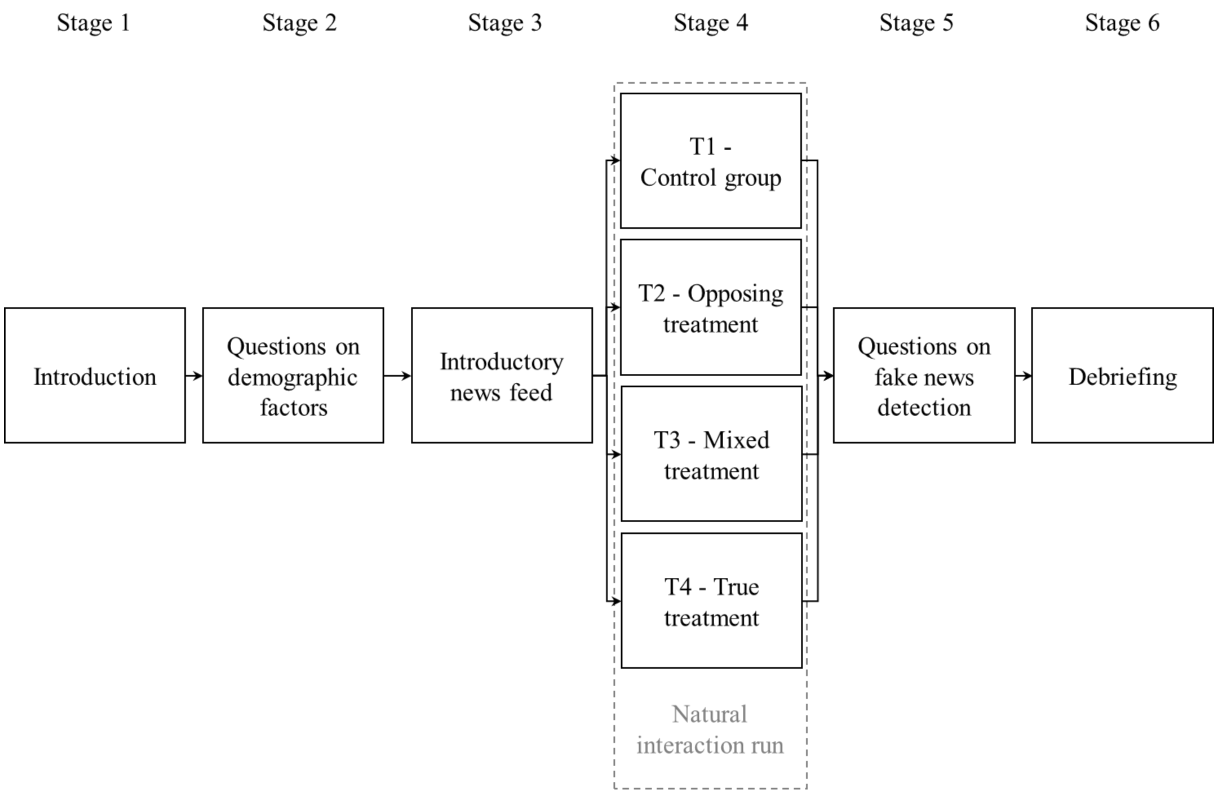


Table 22: Procedure of the online experiment

Appendix B.2 Headlines of Main and Related Articles

Main Headline	*True Related Articles*	*Fake Related Articles*
• *Trump wants military parade for military forces in Washington*	• US President Trump wants military parade* • Trump's scary military obsession* • Trump invests record amount in US military • Why Trump is a danger for peace	• Peace on earth* • Trump wants to cut military spending* • Trump demands: no more money for the military • Trump criticizes Germany for high military spending
• *From abundance to scarcity – crude oil and gas vanish*	• Oil and gas will be scarce soon* • The world without crude oil* • Crude oil and gas won't last another 20 years • When the oil is becoming scarce	• The crude oil lie – it is not actually that scarce* • The myth of finite crude oil* • No sign of scarcity: the secret of crude oil • The energy fallacy: the inexhaustibility of oil and gas
• *The report confirms: systematic doping in Russia*	• Systematic doping: Olympia without Russia* • IOC takes measures against Russia* • Replacement of urine samples detected at another Russian sportsman • Letter to WADA: Russia admits doping for the first time	• Russia and "state doping": the staged scandal* • McLaren report: no proof for Russian doping* • sportsmen in the interview: doping agitation against Russia • No exchange of doping samples: Russia provides proof

• *Merkel hopes for 12 million immigrants*	• CDU and CSU want to limit immigration with seven measures* • Merkel wants to stop immigrants in Africa* • Merkel's big commitment to fighting migration causes • CDU and CSU agree on migration	• Merkel's governmental declaration: mass immigration continues indefinitely* • Merkel wants to advertise for mass immigration to Germany in Africa*
• *Jackie Chan dead: actor dies during filming*	• "I am alive" – Jackie Chan discounts own death* • Jackie Chan is alive! * • After death rumor: Jackie Chan shows sign of life • Jackie Chan: I am not dead	• News: Jackie Chan dead* • Jackie Chan died during fil shooting for RushHour 4? *
• *NASA manipulated climate data: A new Climategate?*	• Conspiracy theory on climate data* • NASA corrects climate data* • Wrong accusations against climate data correction • Correction of climate data is accurate	• Climate data is being forged: the German lie* • Nasa changes climate data retrospectively – why? *

Table 23: Headlines of main and related articles

Note: In T1, no related articles are needed; In T2, four fake related articles are needed for true main articles and four true related articles for fake main articles; In T3, two true and two fake related articles are needed for both, true and fake main articles (marked with *); In T4, four true related articles are needed for both, true and fake main articles

Appendix C Improving Users' Fake News Reporting in Social Media – Two Online Experiments

Appendix C.1 Detailed Sample Description of Study 1

a) Sample Description Study 1

Personal information

Number of participants	320			
Number of participants per treatment	Control	Injunctive SN	Descriptive SN	Combined
	78	83	77	82
Occupation	Students		Employed	
	294		26	
Educational level	General qualification for university entrance	Bachelor's degree	Master's degree	Others
	216	66	17	21
Age	Mean		Standard deviation	
	23.2		4.16	
Gender	Male		Female	
	96		224	

Usage of social media

Social media usage	Several times a day		Once or several times a week	
	253		67	
Trust in news available on social media	No trust and little trust	Rather little trust	Rather a lot of trust	A lot of trust and a great deal of trust

	59	135	101	25	
Content generation compared to their social environment	No content at all	Very little content and little content	Rather little content	Rather a lot of content	A lot of content and a great deal of content
	26	201	58	28	7
Press the like button per day	Mean		Standard deviation		
	4.70		7.15		
Post comment per day	Mean		Standard deviation		
	0.91		1.96		
Share of posts per day	Mean		Standard deviation		
	0.26		1.38		
Report of posts per day	Mean		Standard deviation		
	0.12		0.51		

Table 24: Detailed sample description of Study 1

b) Sample Description of Study 2

Personal information

Number of participants	159			
Number of participants in each treatment	Control	Positive Descriptive SN	Negative Descriptive SN	
	52	55	52	
Occupation	Students		Employed	
	118		33	
Educational level	General qualification for university entrance	University degree	Others	
	61	83	15	
Age	Mean		Standard deviation	
	25.6		5.18	
Gender	Male		Female	
	64		95	
Usage of social media				
Social media usage	Several times a day	Once or several times a week		
	126	33		
Trust in news available on social media	No trust and little trust	Rather little trust	Rather a lot of trust	A lot of trust and a great deal of trust
	44	60	42	13

Content generation compared to their social environment	No content at all	Very little content and little content	Rather little content	Rather a lot of content	A lot of content and a great deal of content
	16	81	41	18	3
Liked posts in comparison to the social environment	None at all	Very few and few	Rather few	Rather many	Many and a great many
	11	55	58	26	9
Commented posts in comparison to the social environment	None at all	Very few and few	Rather few	Rather many	Many and a great many
	43	73	28	12	3
Shared posts in comparison to the social environment	None at all	Very few and few	Rather few	Rather many	Many and a great many
	69	58	20	7	5
Reported posts in comparison to the social environment	None at all	Very few and few	Rather few	Rather many	Many and a great many
	106	31	9	11	2

Table 25: Detailed sample description of Study 2

Appendix C.2 Stimuli

We focus on a mobile device representation because most users access Facebook via mobile devices (Facebook 2016). Nevertheless, the participants can participate in our experiment via desktops or mobile devices. To increase comparability, we, therefore, apply two different views for the news feed. The mobile view shows the news feed about the full size of the browser window. The desktop view first displays the image of a smartphone, which then contains the full-screen size news feed identical to the mobile view.

We further extend the news feed by adding injunctive and descriptive SN messages with different wordings.

a) Posts of Study 1

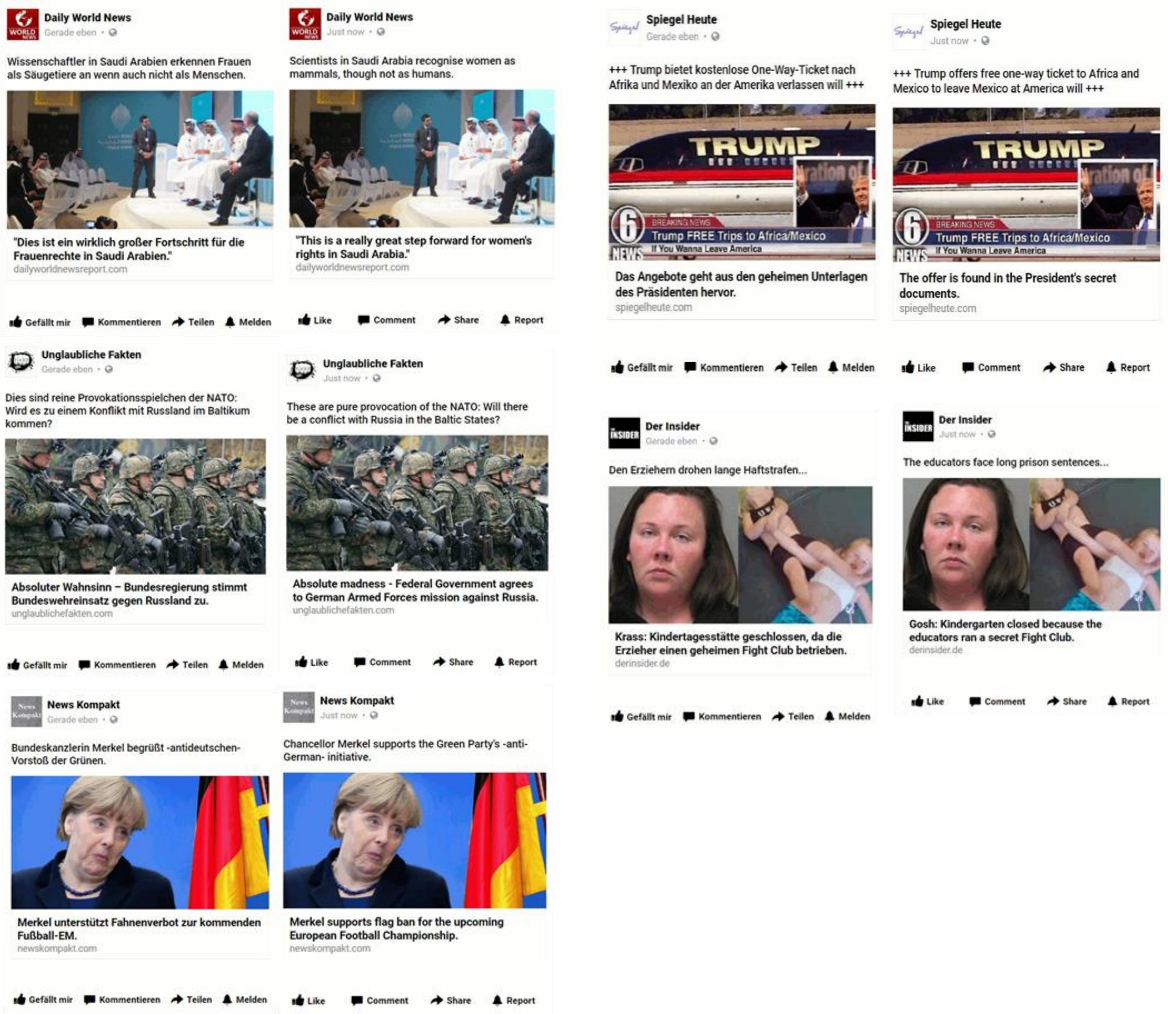


Figure 23: Fake news posts: original German posts (left) & translated English posts (right)

Note: The German (and therefore also the English) text is deliberately wrongly worded, as this property is not unusual for fake news

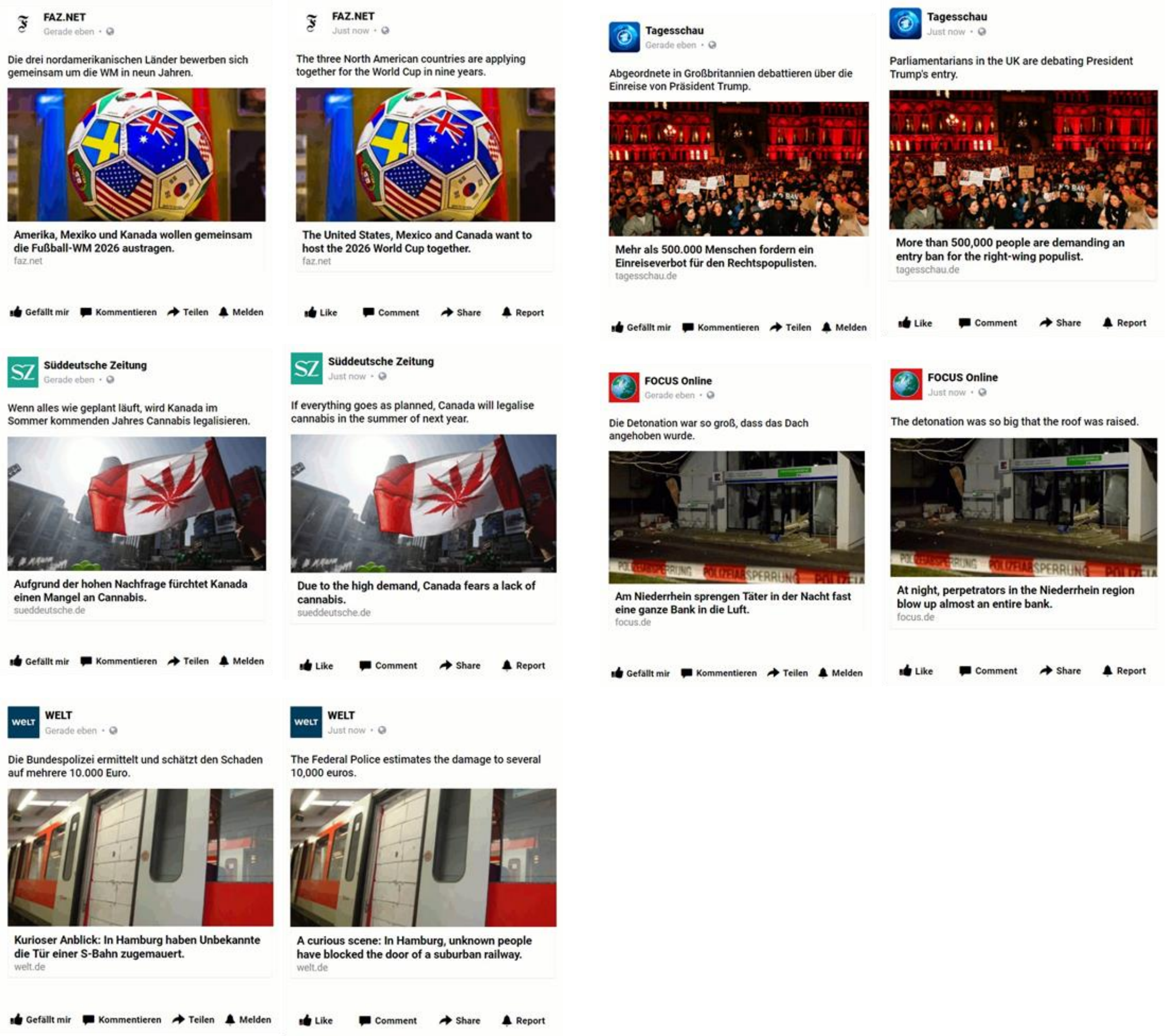


Figure 24: Real posts, original German posts (left) & translated English posts (right)

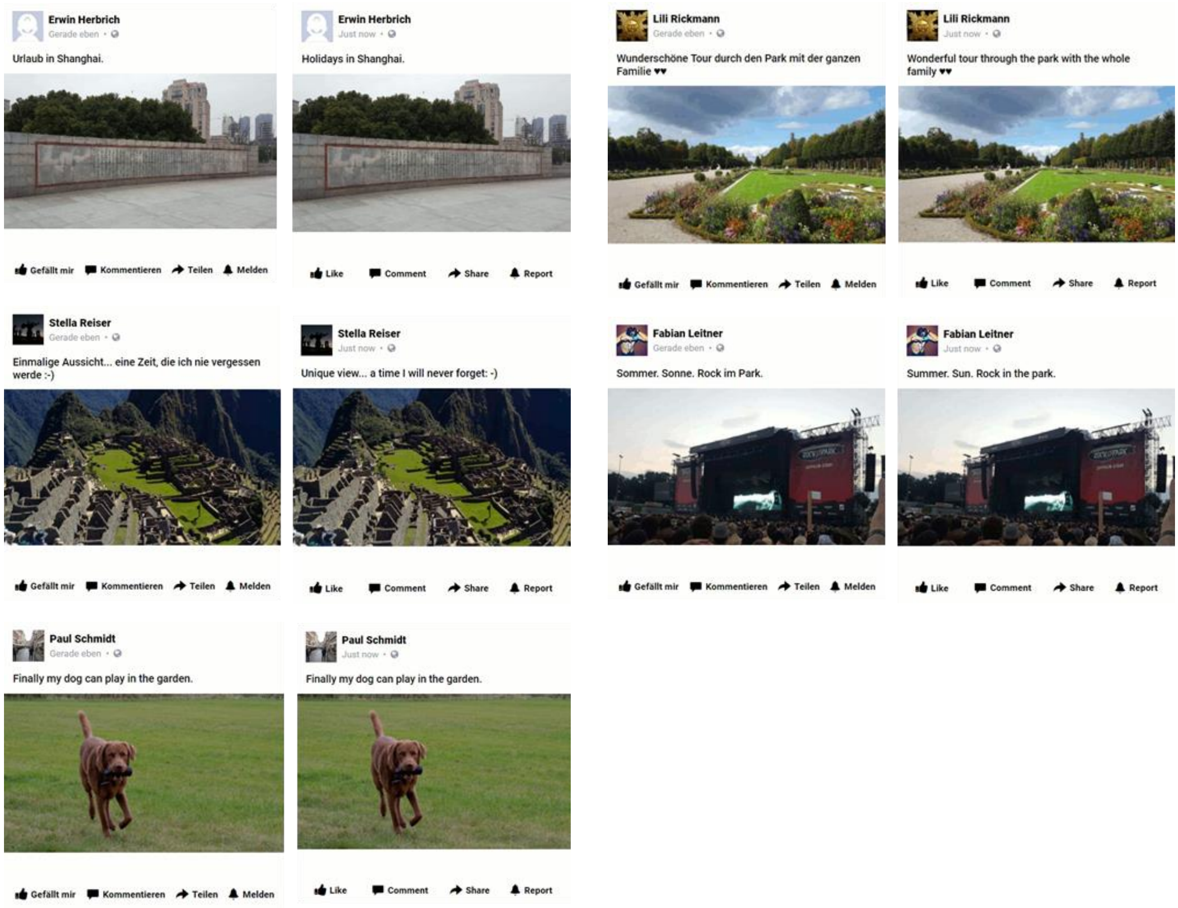


Figure 25: Neutral posts, original German posts (left) vs. translated English posts (right)

b) Posts of Study 2

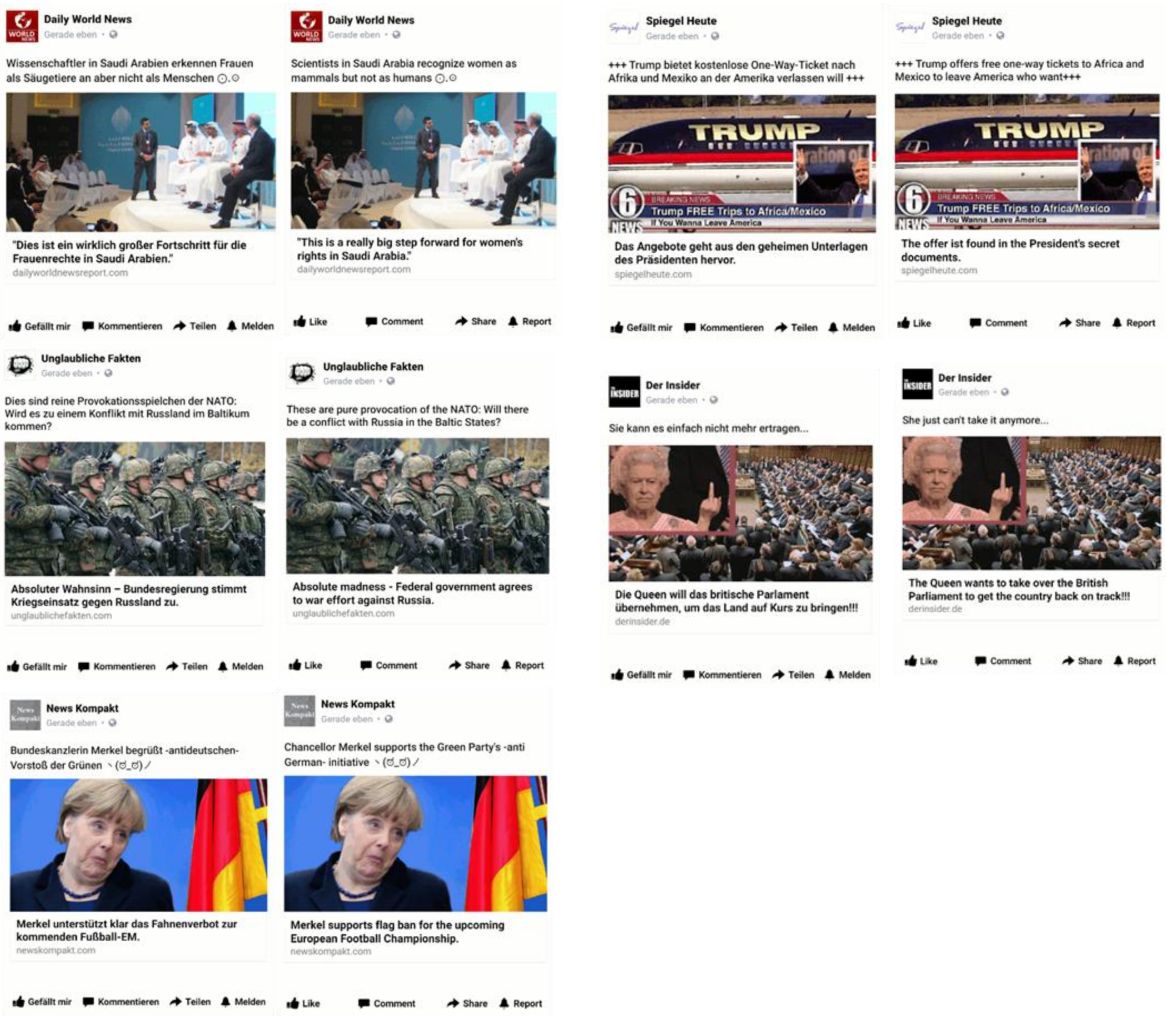


Figure 26: Fake news posts: original German posts (left) & translated English posts (right)

Note: The German (and therefore also the English) text is deliberately wrongly worded, as this property is not unusual for fake news

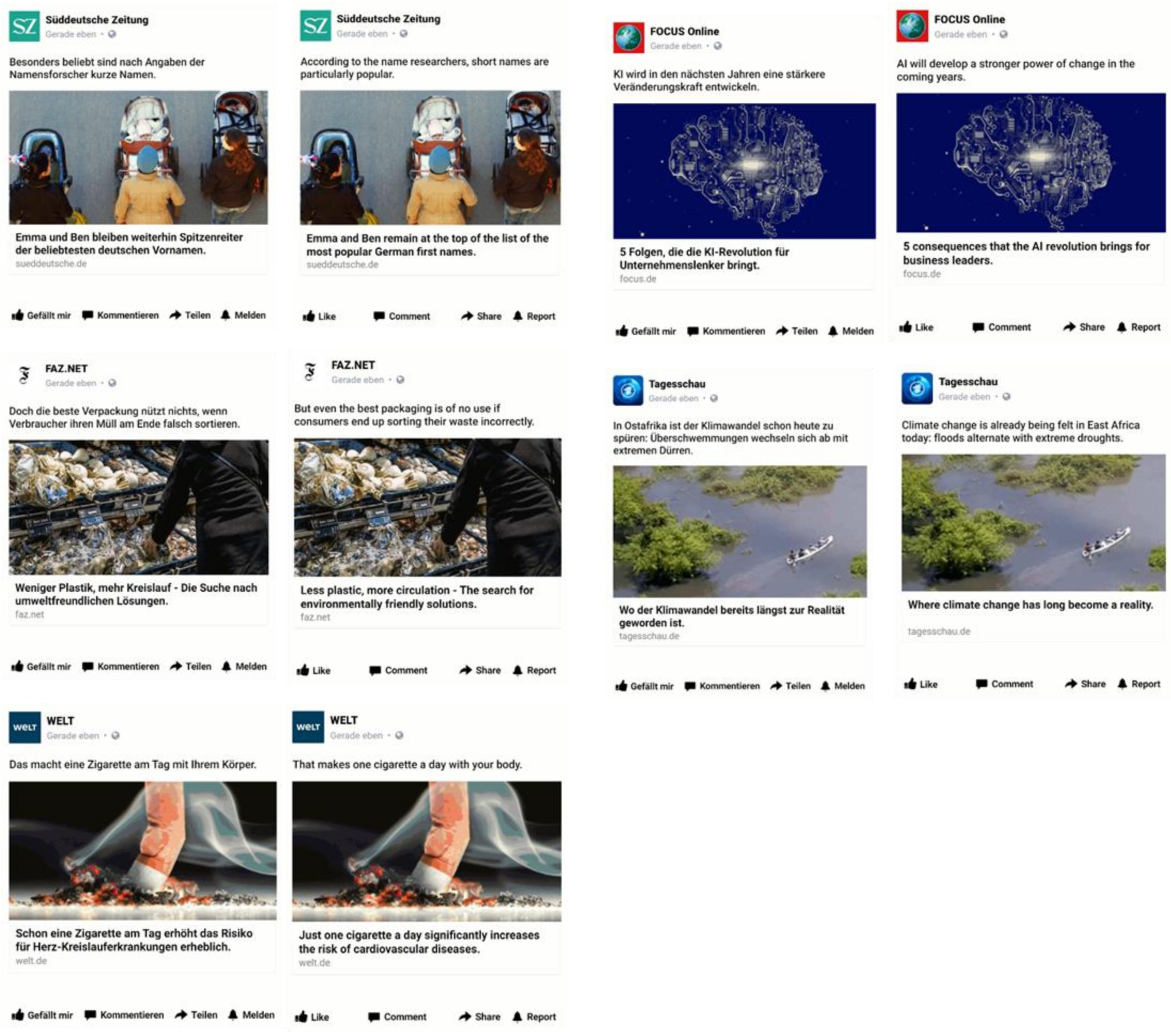


Figure 27: Real posts, original German posts (left) & translated English posts (right)

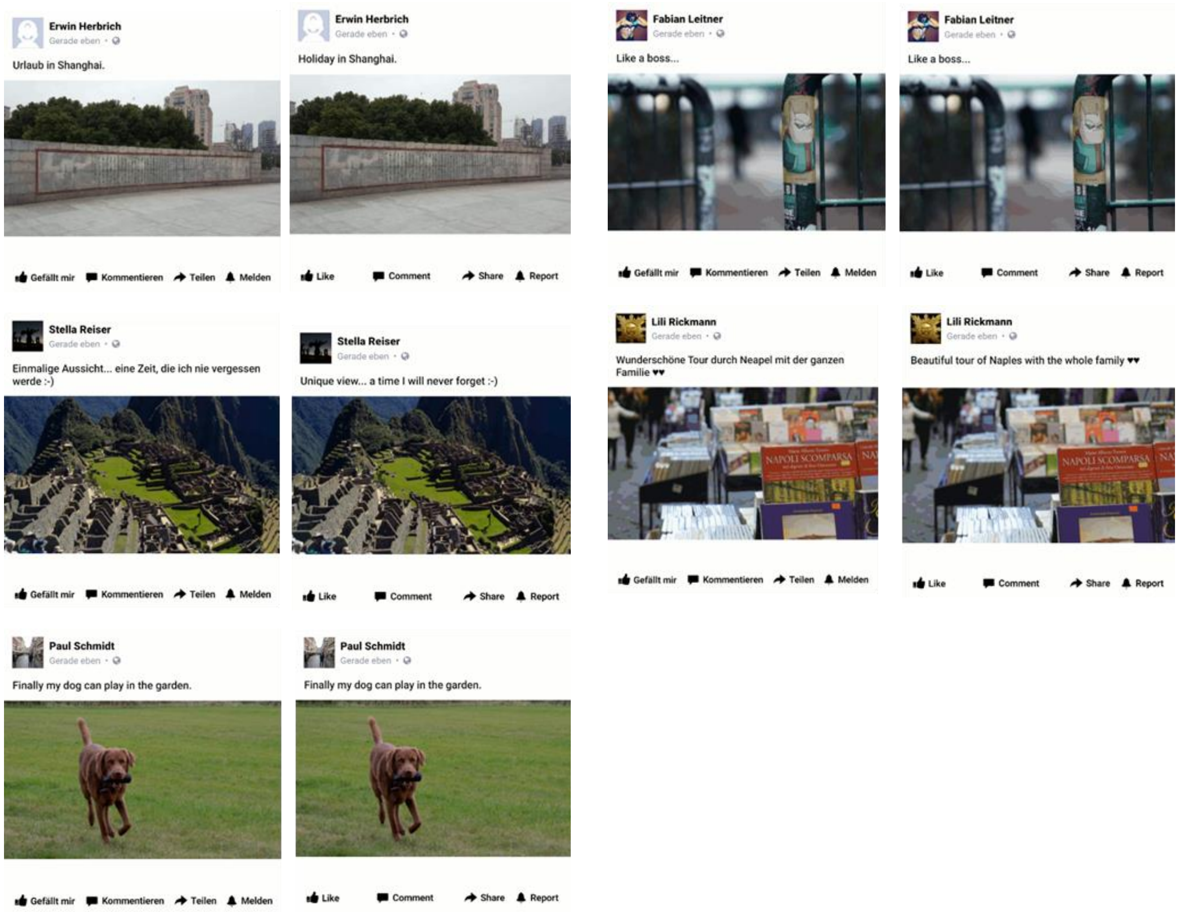


Figure 28: Neutral posts, original German posts (left) vs. translated English posts (right)

Appendix C.3 Results Regarding Structural Equality of Treatments

a) Results of Study 1

	ANOVA		Levene-Test	
Sample property	**F-value**	**p-Value**	**F-Value**	**p-Value**
Age	1.581	0.194	0.595	0.619
Frequency of social media usage	0.059	0.981	0.059	0.981
Trust in news from social media	0.495	0.686	1.010	0.388

	χ^2-test	
Sample property	χ^2	**p-Value**
Gender	4.100	0.231
Occupation	10.262	0.619
Education	23.863	0.470

Table 26: Comparison of the characteristics of the participants in the different treatments on structural equality

Note: + p < 0.1, * p < .05, ** p < .01, *** p < 0.001; n = 320

b) Results of Study 2

	ANOVA		Levene-Test	
Sample property	**F-value**	**p-Value**	**F-Value**	**p-Value**
Age	0.581	0.561	1.980	0.141
Frequency of social media usage	0.102	0.903	0.102	0.903
Trust in news from social media	0.214	0.807	0.125	0.883
	χ^2-test			
Sample property	**χ^2**	**p-Value**		
Gender	2.619	0.283		
Occupation	6.878	0.841		
Education	12.191	0.264		

Table 27: Comparison of the characteristics of the participants in the different treatments on structural equality

Note: + p < 0.1, * p < .05, ** p < .01, *** p < 0.001; n = 159

Appendix C.4 Development and Validation of the News Feed

a) Results of Study 1

Materials: The focus of our research is the reporting of fake news. To consider this effect as isolated as possible, we design the fake news to be as conspicuous as possible. Our goal is to make it rather easy for the participants to identify the fake news to reduce the effect of uncertainty when identifying fake news. However, we deliberately do not want to tell the participants which news posts are fake news because such an implementation does not correspond to the real world and also does the act of reporting unnecessary. A substantial limitation of the external validity of our results would be the consequence of clearly identifying fake news as such. Therefore, we modified the fake news posts based on common fake news characteristics such as spelling mistakes, flashy formatting, or obviously altered images, to make it easier to identify them. Every fake news story is assigned to a non-existent news source to reduce authenticity further. Besides, the fake news headlines were selected, which have spread in the German-speaking area, since the study took place in Germany.

Besides the fake news, the news feed includes five real news posts. These contain no characteristics of fake news and originate from authentic German news sources. Additionally, headlines were chosen, which have spread in the media in German-speaking countries. However, headlines were selected that cannot be identified as real news at first glance. This selection also increases the external validity of our results, as the media landscape is not always clear. At the same time, this allows us to observe how the use of SN affects the reporting behavior of real news.

Overall, especially in comparison to the fake news, the real news can be identified by these characteristics. Besides the news post, the news feed also contains five neutral posts of imaginary friends to create an experimental environment that is as realistic as possible and thus further increase the external validity of the results.

Randomization: The following parts of the experiment were random:

- Assignment of the participants to the treatments
- Display order of the 15 posts by a participant
- Selection of the five fake or real news that is marked with a descriptive SN in treatments where the descriptive SN was used (neutral posts were not marked)

- The strength (equivalent to the number of reports) of a descriptive SN to a post in treatments where the descriptive SN was used
- The strength of the descriptive SN themselves was randomly varied by 10% compared to the base value.

Pre-Test: We conducted a pre-test. The participants of the pre-test (n = 22) were recruited via Facebook, but attention was paid to use different Facebook groups than for the main study to separate the participants of the pre-test from the participants of the studies. The pre-test suggested feasibility and acceptance of the length of the news feed. Based on the pre-test, we adjusted the fake news stories so that they are more easily identifiable by the participants. Further, we examined and adjusted the strength of the descriptive SN to ensure that they fit realistically into the news feed. Our first design included a wider interval for the strength of descriptive SN (1; 10; 100; 1,000; 10,000). However, feedback from the pre-test showed that a norm of 1 is too weak to convince participants that other people report fake news. At the same time, a descriptive SN of 10,000 was found to be too strong. Therefore, we have increased the weakest descriptive SN to 5 and reduced the strongest descriptive SN to 3,125.

Ensuring Response Quality: It was ensured for each participant that the study had been fully processed to ensure the manipulation validity of our studies. Three hundred forty-nine participants have completed the survey. To further improve the quality of our data, we have taken further measures. First, tracking points were implemented into the news feed to identify whether a participant has seen all posts in the news feed. Only participants who had read the news feed completely were considered. Second, we have removed all participants who use social media less than once a week. Our final sample for Study 1 contains 320 participants.

Experiment Validation: To validate our approach, we consider the results of the fake assessment run. Although the fake assessment run should be used with caution (due to the repeated presentation), it is suitable as an indicator to identify which posts the participants considered as fake. Table 28: Number of reports per post in the fake assessment run of Study 1Table 28 summarizes the number of reports for the posts in the fake detection run. The results show that, on average, 79% of the participants identified fake news as such. In contrast, only 20% of the participants reported real news as fake news.

	News Content	Number of reports	Fake detection rate
Fake News	Merkel Flag Ban	273	85.31%
	NATO Mission	262	81.88%
	One-way Ticket	259	80.94%
	Woman Rights	250	78.13%
	Kindergarten Fight Club	219	68.44%
Real News	Legal Cannabis	94	29.38%
	Host World Cup	94	29.38%
	Trump Entry Ban	57	17.81%
	Blocked Wagon	50	15.63%
	Blown Up bank	25	7.81%

Table 28: Number of reports per post in the fake assessment run of Study 1

The following Table 29 compares fake news and real news by counting how often the fake news and real news were reported and how often they were not reported. Fisher's exact test shows a highly significant empirical difference (p-value < 0.001). These results are in line with our expectations since both the fake news and the real news were identified as such by the participants.

	Reported	Not reported
Fake News	1,263	337
Real News	320	1,280

Table 29: Overall number of reports in the fake assessment run of Study 1

b) Results of Study 2

Materials: For Study 2, we used almost the same materials as for Study 1. However, we modified four of the five original fake news to further decrease the level of authenticity. Emoticons (such as ‘ヽ(ʘ_ʘ)ノ’ and ‘⨀.⊙’) were added to the news posts about women's rights in Saudi Arabia and the ban on flags at the European Championship to give participants the additional feeling that the news source is not neutral, but that the author wants to give a special feeling and opinion, as it is often the case with fake news. At the same time, we changed the term 'German Armed Forces Mission' to 'War mission' in the fake news post about the NATO mission. We intended to differentiate the statement even more from the real world and thus make it less credible.

Additionally, we replaced one fake news completely ("Kindergarten Fight Club"), because it was recognized much less by the participants than the others. We assume that this fake news was harder to identify because it does not relate to events with which the participants frequently come into contact.

The real news was also revised. Since the real news stories were outdated at the time of Study 2, we replaced them with more recent ones. We have also adjusted the neutral posts, as these explicitly referred to events in the summer (the time of Study 1). Since Study 2 took place in winter, we have exchanged these posts to make our news feed more realistic.

Randomization: As in Study 1

Pre-Test: Since our material and procedure had already been validated by the pre-test and Study 1, we did not conduct another pre-test for Study 2.

Ensuring Response Quality: We have taken the measures described above for Study 2 with two additions. First, we asked the participants whether they had already participated in Study 1 and removed all participants who confirmed this. Second, we added a control item to each of our scales to record the motivational factors for reporting. This control item corresponds to the inverted statement of another, which is why a response at the same end of the scale is inconsistent. Table 30 presents the items described.

Target construct	*Test item*	*Control item*
Motivation to report fake news	I would like to avoid negative consequences that arise from the distribution of fake news	I am not interested in the negative consequences that arise from the distribution of fake news
Motivation to not report fake news	I've never seen fake news	I often come in contact with fake News

Table 30: Summary of control items in Study 2

We have removed all participants who answered the control item inconsistently. Therefore, our final data set for Study 2 consists of 159 participants.

Experiment Validation: To validate our approach, we consider the results of the updated fake assessment run. Table 31 summarizes the number of reports for the posts in the fake detection run. The results show that, on average, 87% of the participants identified fake news as such. In contrast, only 3% of the participants reported real news as fake news.

	News Content	*Rated as fake*	*Fake detection rate*
Fake News	British Parliament	155	97.48%
	Merkel Flag Ban	145	91.19%
	NATO Mission	142	89.31%
	Woman Rights	128	80.50%
	One-way Ticket	120	75.47%
Real News	AI Influence	7	4.40%
	Climate Change	7	4.40%
	Effects of Smoking	4	2.51%
	Reuse of Shopping Bags	3	1.89%
	First Name Raking	3	1.89%

Table 31: Number of news stories rated as fake news in Study 2

Table 32 compares fake news and real news by counting how often the fake and real news were rated as fake and how often they were not rated as fake. Fisher's exact test shows a highly significant empirical difference (p-value < 0.001). These results are in line with our expectations since we deliberately made the difference between fake and real news more extreme for Study 2.

	Rated as fake	*Not rated as fake*
Fake News	690	105
Real News	24	771

Table 32: Overall number of news stories rated as fake news in Study 2

Appendix C.5 List of Scales

a) *List of Scales for Study 1*

Construct	*Question*	*Scale type*	*Response options*	*Source*
Social media as a source of news	How often do you use social media to obtain information on current events, public affairs, and political issues?	9-point (ordinal) Likert scale	Never; Less than once a week; Once a week; Twice a week; Three times a week; Four times a week; Five times a week; Six times a week; Daily	Based on the scale for measuring the frequency of the use of different media as a news source by (Gil de Zúñiga et al. 2012).
	How much trust do you have in information about current events, public affairs, political issues that comes from social media?	6-point (ordinal) Likert scale	No trust at all; Little trust; Rather little trust; Rather a lot of trust; A lot of trust; A great deal of trust	
Indicators for Fake News	What features did you use in the survey, and do you generally use to identify a post as fake news?	Multiple Choice with optional text input	Source/Creator; Content of the text; Spelling mistake in the text; Striking formatting of the text; Picture in the post; Number of times a mail has already been reported; others	Self-developed

Usage of social media	How often do you use social media in your everyday life?	6-point (ordinal) Likert scale	Never; Less than once a week; Once a week; Several times a week; Once a day; Several times a day	
	How much content do you generate in social media (e.g., by posting, liking, commenting, sharing, etc.) compared to people from your personal environment?	7-point (ordinal) Likert scale	No content at all; Very little content; Little content; Rather little content; Rather a lot of content; A lot of content; A great deal of content	
	How many posts do you like per day?	Numerical input	Integer value equal to or greater than 0	Self-developed
	How many posts do you comment per day?	Numerical input	Integer value equal to or greater than 0	
	How many posts do you share per day?	Numerical input	Integer value equal to or greater than 0	
	How many posts do you report per day?	Numerical input	Integer value equal to or greater than 0	
Automatic vs. mindful thought processing	I don't like to have to do a lot of thinking.	5-level semantic differential	Min: Strongly disagree Max: Strongly agree	
	I try to avoid situations that require	5-level semantic	Min: Strongly disagree	

	thinking in depth about something.	differential	Max: Strongly agree	Self-translated scale to measure the degree to which one engages in mindful versus automatic thought-processing of news by (Maksl et al. 2015).
	I prefer to do something that challenges my thinking abilities rather than something that requires little thought.	5-level semantic differential	Min: Strongly disagree Max: Strongly agree	
	I prefer complex to simple problems.	5-level semantic differential	Min: Strongly disagree Max: Strongly agree	
	Thinking hard and for a long time about something gives me little satisfaction.	5-level semantic differential	Min: Strongly disagree Max: Strongly agree	
Cognitive reflection test	A racket and a ball cost a total of 1.10 Euros. The racket costs 1.00 Euro more than the ball. How much does the ball cost?	Numerical input	Decimal number equal to or greater than 0	Self-translated scale to measure reflective reasoning by (Frederick 2005).
	Five machines take 5 minutes to produce five products. How many minutes does it take 100 machines to produce 100 products?	Numerical input	Integer value equal to or greater than 0	

	Lilies grow in a lake. Every day, the area that covers the lilies on the lake doubles. If it takes 48 days until the lilies cover the whole lake, how long would it take until the lilies cover half of the lake?	Numerical input	Integer value equal to or greater than 0	
Gender	What is your gender?	Single Choice	female; male	
Age	How old are you?	Numerical input	Integer value equal to or greater than 0	
Educational level	What is your educational level? Please select the highest level of education you have achieved so far.	Single Choice with optional text input	Leaving school without graduation; Still in school; junior high school (original: “Volks-, Hauptschulabschluss”, “Mittlere Reife”, “Realschule”) or equivalent degree; Completed apprenticeship; Senior High School (Original: “(Fach-) Abitur”); Bachelor’s degree; Master’s degree; Diploma; Doctorate / postdoctoral	Based on the standard demographic questions of the Sosci Survey tool (www.soscisurvey.de)

			qualification; Different degree [free text option]
Occupation	Which of the following categories best describes your employment status?	Single Choice with optional text input	Pupil; In an apprenticeship; Student; Salaried employee; Civil servant; Independent; Unemployed; Others [free text option]

Table 33: Description of all scales used in Study 1

b) List of Scales for Study 2

Construct	*Question*	*Scale type*	*Response options*	*Source*
Social media as a source of news	How often do you use social media to obtain information on current events, public affairs, and political issues?	9-point (ordinal) Likert scale	Never; Less than once a week; Once a week; Twice a week; Three times a week; Four times a week; Five times a week; Six times a week; Daily	Based on the scale for measuring the frequency of the use of different media as a news source by (Gil de Zúñiga et al. 2012).
	How much trust do you have in information about current events, public affairs, political issues that comes from social media?	6-point (ordinal) Likert scale	No trust at all; Little trust; Rather little trust; Rather a lot of trust; A lot of trust; A great deal of trust	
Assessment of news posts	Please rate whether the post shown above is a true news story or fake news:	5-point (ordinal) Likert scale	True news story; Rather a true news story; I can't tell; Rather fake news; Fake news	Self-developed
Indicators for Fake News	What features did you use in the survey, and do you generally use to identify a post as fake news?	Multiple Choice with optional text input	Source/Creator; Content of the text; Spelling mistake in the text; Striking formatting of the text; Picture in the post; Number of times a mail has already been reported; others	Self-developed

Reasons to report fake news	I would like to improve the living conditions for myself and other people	5-level semantic differential	Min: Strongly disagree Max: Strongly agree	Self-developed
	It is important to me to help other people to form their opinion based on true facts	5-level semantic differential	Min: Strongly disagree Max: Strongly agree	
	I would like to avoid negative consequences that result from the dissemination of fake news	5-level semantic differential	Min: Strongly disagree Max: Strongly agree	
	Fake News cause an unpleasant feeling in me	5-level semantic differential	Min: Strongly disagree Max: Strongly agree	
	I would like to improve the quality of the social media platform I use	5-level semantic differential	Min: Strongly disagree Max: Strongly agree	
	I expect a material or financial incentive	5-level semantic differential	Min: Strongly disagree Max: Strongly agree	
	A correct news landscape is important to me	5-level semantic differential	Min: Strongly disagree Max: Strongly agree	

	I want to share my knowledge with others	5-level semantic differential	Min: Strongly disagree Max: Strongly agree
	I hope for appreciation from my social environment	5-level semantic differential	Min: Strongly disagree Max: Strongly agree
	I am not interested in the negative consequences that arise from the distribution of fake news	5-level semantic differential	Min: Strongly disagree Max: Strongly agree
	I do not want that the relevance of the topic will be reduced by the dissemination of fake news	5-level semantic differential	Min: Strongly disagree Max: Strongly agree
Hurdles to report fake news	I am not interested in public opinion	5-level semantic differential	Min: Strongly disagree Max: Strongly agree
	The procedure of reporting is too complex for me	5-level semantic differential	Min: Strongly disagree Max: Strongly agree
	I've never seen fake news	5-level semantic	Min: Strongly disagree Max: Strongly agree

	differen-tial	
I do not believe that the act of reporting has any effect or counteracts the dis-semination of fake news	5-level semantic differen-tial	Min: Strongly disa-gree Max: Strongly agree
I don't know how to report fake news	5-level semantic differen-tial	Min: Strongly disa-gree Max: Strongly agree
Reporting fake news offers me no material or financial incentive	5-level semantic differen-tial	Min: Strongly disa-gree Max: Strongly agree
Fake news is subject to the freedom of speech	5-level semantic differen-tial	Min: Strongly disa-gree Max: Strongly agree
Fake news serves for entertainment	5-level semantic differen-tial	Min: Strongly disa-gree Max: Strongly agree
I often come in con-tact with fake news	5-level semantic differen-tial	Min: Strongly disa-gree Max: Strongly agree

	I am only a consumer of news and do not actively participate in the public discourse	5-level semantic differential	Min: Strongly disagree Max: Strongly agree	
Usage of social media	How often do you use social media in your everyday life?	6-point (ordinal) Likert scale	Never; Less than once a week; Once a week; Several times a week; Once a day; Several times a day	Self-developed
	How much content do you generate in social media (e.g., by posting, liking, commenting, sharing, etc.) compared to people from your personal environment?	7-point (ordinal) Likert scale	No content at all; Very little content; Little content; Rather little content; Rather a lot of content; A lot of content; A great deal of content	
	How many posts have you liked in the last month compared to people from your personal environment?	7-point (ordinal) Likert scale	None at all; Very few; Few; Rather few; Rather many; Many; A great many	
	How many posts have you commented on in the last month compared to people from your personal environment?	7-point (ordinal) Likert scale	None at all; Very few; Few; Rather few; Rather many; Many; A great many	

	How many posts have you shared in the last month compared to people from your personal environment?	7-point (ordinal) Likert scale	None at all; Very few; Few; Rather few; Rather many; Many; A great many	
	How many posts have you reported in the last month compared to people from your personal environment?	7-point (ordinal) Likert scale	None at all; Very few; Few; Rather few; Rather many; Many; A great many	
Gender	What is your gender?	Single Choice	female; male	
Age	How old are you?	Numerical input	Integer value equal to or greater than 0	
Educational level	What is your educational level? Please select the highest level of education you have achieved so far.	Single Choice with optional text input	Leaving school without graduation; Still in school; junior high school (original: “Volks-, Hauptschulabschluss”, “Mittlere Reife”, “Realschule”) or equivalent degree; Completed apprenticeship; Senior High School (Original: “(Fach-) Abitur”); Bachelor’s	Based on the standard demographic questions of the Sosci Survey tool (www.soscisurvey.de)

			degree; Master's degree; Diploma; Doctorate / postdoctoral qualification; Different degree [free text option]
Occupation	Which of the following categories best describes your employment status?	Single Choice with optional text input	Pupil; In an apprenticeship; Student; Salaried employee; Civil servant; Independent; Unemployed; Others [free text option]

Table 34: Description of all scales used in Study 2

Appendix C.6 Results of the Brant Test for all Ordered Logistics Regression Models

a) Results of Study 1

Base Treatment	*Treatment*	*Coefficient*	*df*	*Probability*
Control	Omnibus	4.205	12	0.979
	Injunctive SN	0.301	4	0.990
	Descriptive SN	1.198	4	0.878
	Combined	2.970	4	0.563
Combined	Omnibus	4.205	12	0.979
	Control	2.970	4	0.563
	Injunctive SN	2.126	4	0.713
	Descriptive SN	0.924	4	0.921

Table 35: Results of the Brant test to confirm the assumption of proportional odds for the ordered logistic regression models of Study 1

Note: + p < 0.1, * p < .05, ** p < .01, *** p < 0.001; n = 320

b) Results of Study 2

Base Treatment	*Treatment*	*Coefficient*	*df*	*Probability*
Control	Omnibus	10.016	8	0.264
	Positive descriptive SN	2.300	4	0.245
	Negative descriptive SN	5.445	4	0.680

Table 36: Results of the Brant test to confirm the assumption of proportional odds for the ordered logistic regression model of Study 2

Note: + p < 0.1, * p < .05, ** p < .01, *** p < 0.001; n = 159

Appendix C.7 Results of the Ordered Logistic Regression with Respect to Real News Reporting of Study 1

Base Treatment	*Treatment*	*Coeffi-cient*	*Standard Error*	*p-Value*		*Odds Ratio*
Control	Injunctive SN	0.847	0.709	0.232		2.333
	Descriptive SN	1.200	0.687	0.080	+	3.320
	Combined	1.482	0.666	0.026	*	4.402
						Nagelkerke's R^2: 0.037
Combined	Control	-1.482	0.666	0.026	*	0.227
	Injunctive SN	-0.635	0.503	0.207		0.530
	Descriptive SN	-0.281	0.472	0.552		0.755
						Nagelkerke's R^2: 0.037

Table 37: Results of the ordered logistic regression to compare the four treatments with respect to real news reporting. In the first model, the control treatment is the baseline, whereas, in the second model, the combined treatment is the baseline

Note: + $p < 0.1$, * $p < .05$, ** $p < .01$, *** $p < 0.001$; n = 320

Appendix C.8 Regression Models on the Strength of Descriptive Social Norms

a) Natural interaction run – fake news posts

Fake posts		*Coefficients*											
		0		*5*		*25*		*125*		*625*		*3125*	
Baselines	*0*	-1.473	***	0.723	*	0.823	*	1.160	***	1.306	***	0.570	
	5			-0.750		0.100		0.437		0.583		-0.153	
	25					-0.651	+	0.337		0.484		-0.252	
	125							-0.314		0.147		-0.589	+
	625									-0.167		-0.736	*
	3125											-0.903	**

Table 38: Coefficients (upper triangle) and intercepts (diagonal) of the logistic regression models with different baselines for fake news posts in the natural interaction run

Note: + p < 0.1, * p < .05, ** p < .01, *** p < 0.001; n = 410

b) Natural interaction run – real news posts

Real posts		*Coefficients*											
		0		*5*		*25*		*125*		*625*		*3125*	
Baselines	*0*	-3.475	***	0.142		0.361		1.047		0.702		0.836	
	5			-3.332	**	0.219		0.905		0.560		0.693	
	25					-3.114	***	0.686		0.341		0.475	
	125							-2.428	***	-0.345		-0.211	
	625									-2.772	***	0.134	
	3125											-2.229	***

Table 39: Coefficients (upper triangle) and intercepts (diagonal) of the logistic regression models with different baselines for the real news posts in the natural interaction run

Note: + p < 0.1, * p < .05, ** p < .01, *** p < 0.001; n = 410

Appendix D Value Stream Modeling and Notation (VSMN) – Developing a Domain-Specific Modeling Language

Appendix D.1 List of modeling approaches

	Method	***Design Objective 1***	***Design Objective 2***	***Design Objective 3***
Artifacts in Lean Management	process standardization	-	-	-
	shop floor management	-	-	-
	idea management	-	-	-
	one-piece flow	-	-	-
	quick changeover	-	-	-
	value stream method (VSM)	-	X (partly)	X (partly)
	U layout	-	-	-
	Kanban	-	-	-
Artifacts in Information Systems	Petri nets	-	X (partly)	X (partly)
	Unified Modelling Language (UML)	-	-	X (partly)
	Business Process Model and Notation (BPMN)	X (partly)	-	X (partly)

Table 40: List of modeling languages/methods in lean management following VDI 2870 (2013) and IS (Object Management Group 2011, 2015; Petri 1962)

Appendix D.2 Design Iterations

No.	***Participants***	***Purpose and outcome***	***Scenario***
1	All authors	The authors discuss the suitability of existing modeling languages for the use case in our research project.	Goods receipt
2	Subgroup of authors	The first design of a VSM extension is designed. The result is the separation of material-oriented and information-oriented process steps, process parameters, the location construct, the event construct, and the construct for digital technologies. Furthermore, arrows for information flow between activities are introduced following *SequenceFlow* notation element of BPMN.	Goods receipt
3	All authors	Discussion of the extensions and introduction of the resting phase element to visualize temporal expansions of buffers in the material flow due to information-oriented activities or unavailable information.	Goods receipt
4	Subgroup of authors	Introduction of a separate symbol for information flows between material-oriented and information-oriented activities. Discussion of gateways as an appropriate construct in value streams.	Goods receipt
5	Subgroup of authors	The authors adjusted the timeline and distinction between material-oriented and information-oriented lead time.	Goods receipt
6	Subgroup of authors	Modeling of the value stream in the scenario of bed logistics. Refinement of the event construct.	Bed cleaning and provision
7	All authors	Discussion of the metamodels, including the correction of logical mistakes.	-

8	Subgroup of authors	Modeling of the value stream in the scenario of material supply at hospital wards. Revision of the information flow element between material-oriented and information-oriented activities following the message flow notation of BPMN.	Material supply at a hospital ward
9	Authors and experts	Evaluation of VSMN through semi-structured interviews and refinement of notation elements (see Appendix D.3 for details).	Goods receipt
10	All authors	The final discussion of VSMN after the evaluation interviews have taken place.	Goods receipt and bed cleaning

Table 41: Overview of design iterations including purpose and outcome, as well as applied demonstration scenarios

Appendix D.3 Expert Interviews and Demonstration Scenario

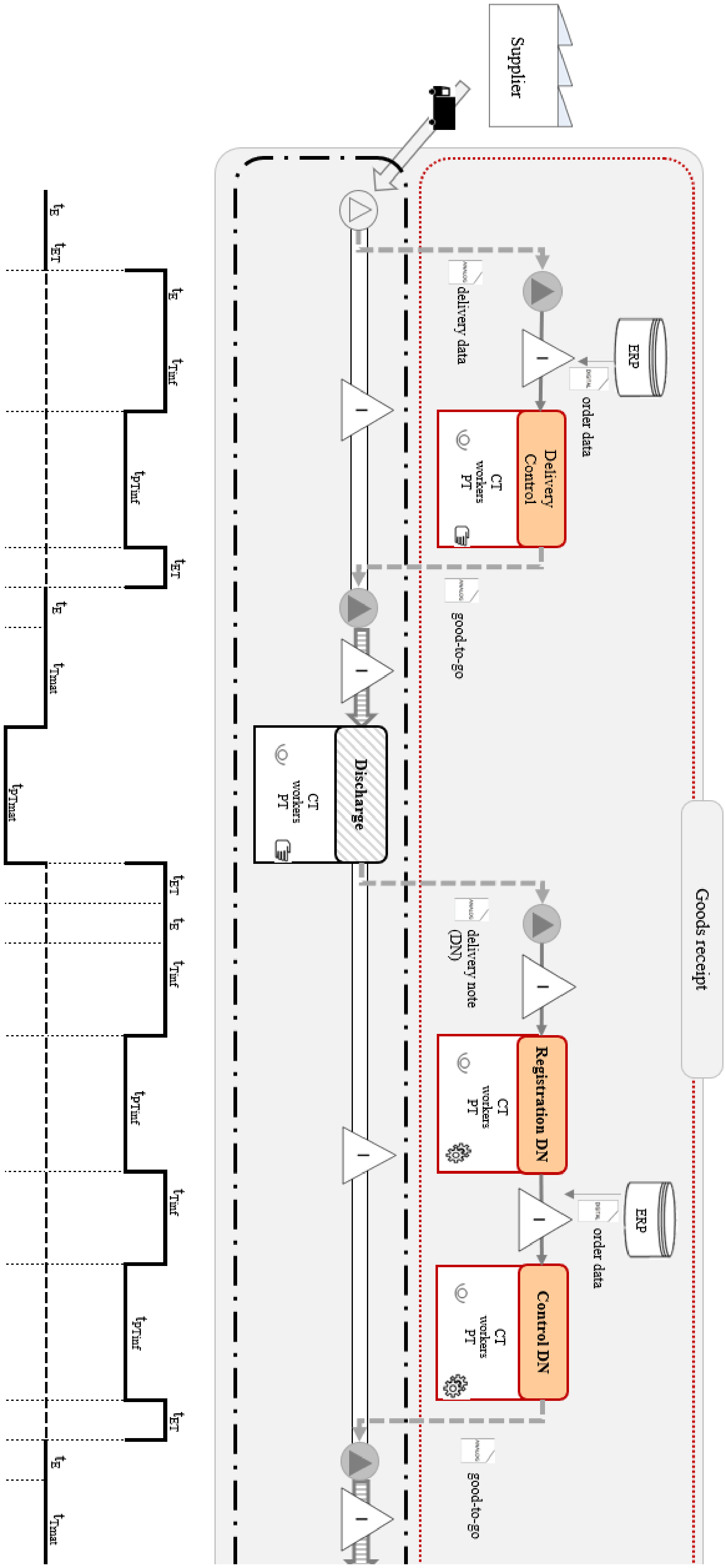
a) Demonstration of VSMN for the Goods Receipt Scenario of a Hospital

VSMN in this demonstration scenario is an earlier version as the feedback of the experts is not incorporated already (Figure 29). The demonstration is a goods receipt scenario, which begins with the transport of the goods from a supplier to the goods entrance (see Chapter 5.2 for the latest version). The physical delivery of the goods triggers the first event in the material flow of the value stream. Information about the order is required to continue the material flow, and a transfer from the material to the information flow occurs. During this, a buffer in the material flow occurs, which is marked with the triangle. The transfer of the information (t_{ET}) about the delivery is time measured, as well as the event (t_E) that triggers the following information activity *delivery control*. The first activity is located in the goods receipt area. It is a manual process with a certain processing time (t_{PTmat}). As long as the respective information-oriented activities have not been executed, the material lies in the subsequent buffer. The resting phase symbol visualizes the time expansion.

The activity *delivery control* creates good-to-go information, which triggers the subsequent activity *discharge*. Subsequently, the information is transferred and causes an event in the material flow. After the activity *discharge* has been completed, analog information in the form of the delivery note is again transferred from material flow to information flow. The transfer of the information about the delivery note from the material flow to the information flow is critical. As long as the delivery note is available but not used, the information stays in the information buffer until the subsequent activity *capture delivery note* is executed. This automatic process converts analog information into digital information. The information is then enriched with order data from the enterprise resource planning system, and the activity *control delivery note* is executed. The activity creates another good-to-go information report, which is required to initiate the push material flow.

A new buffer can emerge in front of the subsequent material-oriented activity *goods inspection*, as long as the goods are not processed immediately. This activity is also a manual activity where the delivered material is inspected and further information created, which is needed by the information-oriented, subsequent activity *posting of goods*. Consequently, another transfer of posting data from material flow to information flow occurs. This data contains analog information on serial numbers or expirations dates, for instance, which the employee captures from the goods. The activity *posting of goods* creates information, which is stored in the enterprise resource planning system. The enterprise resource planning system then enriches the posting information with additional information on the storage location.

The next activity, *creation of a storage order* in the information flow, creates a storage order, which is printed. After this step, all necessary information for the further transport of the material is available. This analog information is transferred (t_{ET}) from the information flow to the material flow and triggers another intermediate event (t_E) in the demonstration process. Between the activity *goods control* and the transport to the next activity *storage*, the material waits for the necessary information. In this time, the material is in a resting phase again, which is illustrated with the respective symbol. After transport, the material is stored in the warehouse. With this last activity *storage* – which is located at the warehouse – the considered demonstration process ends. Accordingly, the sink is missing in this demonstration.



Supplier
Goods receipt
delivery data
ERP
order data
Delivery Control
CT workers PT
good-to-go
Discharge
CT workers PT
delivery note (DN)
Registration DN
CT workers PT
ERP
order data
Control DN
CT workers PT
good-to-go
t_E
t_{ET}
t_{Tinf}
t_{PTinf}
t_{Tmat}
t_{PTmat}

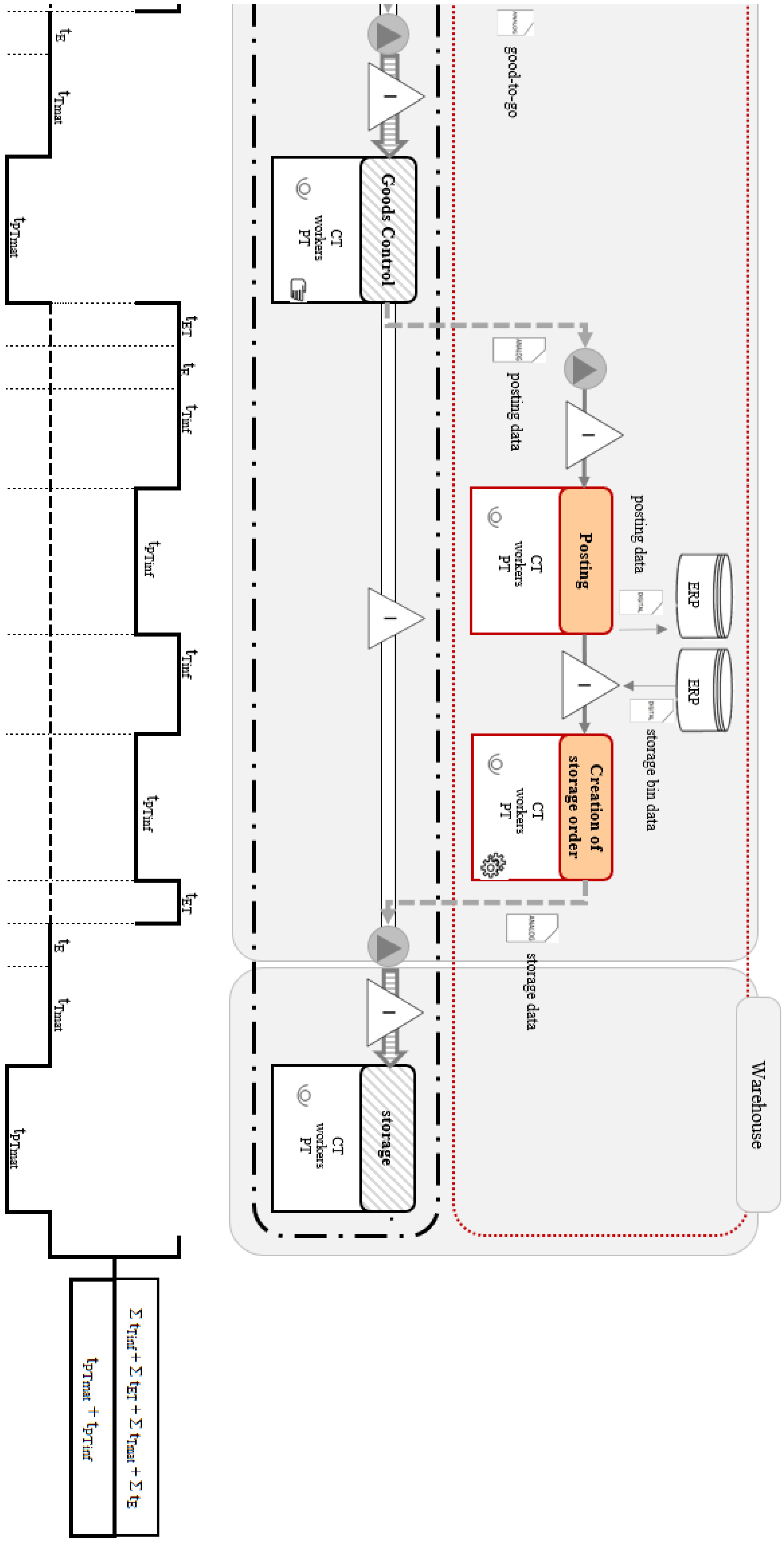


Figure 29: Demonstration scenario, similar as used for expert interviews

Summary of Expert Interviews

Expert	*Relevance*	*Use case in the sector*	*Most relevant feedback on VSMN*
Production I	Today still many breaks in media that control value creation processes. For this reason, many inefficiencies occur.	VSMN allows the analysis of information demand in manufacturing processes.	The resting phase element is appropriate but missing in the meta-model. The separate visualization of material-oriented and information-oriented activities is useful.
Production II	Appropriate information availability allows lean processes not only in logistics but especially in production and service processes.	Transfer of information between material-oriented and information-oriented activities or digital technologies is prone to waste and must be analyzed particularly carefully.	VSMN is suitable for the intended purpose
Production III	Connected processes within the company and across company boundaries are essential in the automotive industry.	Integration on shop floor level to steer production processes and offer analytical services, e.g., predictive maintenance.	*Environment* renamed to *Location* and *Means of Transport* to *Resource*. Revision of activity symbol colors to allow black and white visualization.
Service I	Due to the well-known basis of the VSM, the VSMN can easily be applied in consulting	Depending on the focus, different areas within a company can be analyzed.	Separate illustration of information flow and material flow is introduced.

			Modify the color illustration of the information flow, because red is generally understood as negative.
Service II	Lack of transparency of stocks and materials is a major problem and has many inefficiencies	Providing comprehensive logistics services through-out the hospital	VSMN is well suited to compare process variants. VSMN allows seeing which process steps, especially in the information flow, can be moved in advance or combined in order to reduce the processing time. VSMN is too complex to communicate the processes to all employees. Simplification of VSMN in some symbols to achieve less complexity. 1) only one symbol for digital and analog data. 2) a unified symbol for buffers. 3) a single symbol for events.

Table 42: Summary of insights from expert interviews