Learning with Gamified Information Systems

The Role of Gamification Design Elements, Contexts, and Ethical Implications

DISSERTATION

zur Erlangung des Grades eines Doktors

rerum politicarum

der Fakultät III – Wirtschaftswissenschaften,
Wirtschaftsinformatik und Wirtschaftsrecht
der Universität Siegen

vorgelegt von

Rahel-Katharina Jahn

Erstgutachter: Univ.-Prof. Dr. Dr. Björn Niehaves

Zweitgutachter: Univ.-Prof. Dr. Tim Klucken

Dekan: Univ.-Prof. Dr. Marc Hassenzahl

Datum der Disputation: 02. März 2022

Acknowledgements

I would like to thank the following people for their support. First, my supervisor, Björn Niehaves, for his invaluable advice, continuous support, and assistance at every stage of my research projects. The mentorship he provided helped me to gain profound academic as well as professional knowledge and skills, especially in the discipline of Information Systems.

Second, I would like to express my sincere thanks to my colleagues Oliver Heger, Bastian Kordyaka, and Anika Nissen, with whom I could develop, discuss, and conduct a range of research projects. Additionally, I also thank all my colleagues at University of Siegen and Chemnitz University of Technology for the wonderful interdisciplinary experiences. Specifically, I would like to mention Tanja Joan Eiler, Henrik Freude, Jacqueline Gräf, Hans Christian Klein, Michael Klesel, Kristina Lemmer, Alla Machulska, Marius Müller, Frederike Marie Oschinsky, Oliver Rehren, Rewa Tamboli and Andreas Weigel.

Moreover, I would like to thank my partner, friends, and family for their ongoing support. Finally, I would like to thank all participants of my studies without whom this dissertation would not have been possible.

Siegen, March 2022 Rahel-Katharina Jahn

Contents

List of Tables

List of Figures

Glossary

AAT	Approach Avoidance Task
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
AR	Augmented Reality
AT	Affective Technology
AVE	Average Variance Extracted
BYOD	Bring Your Own Device
CA	Cognitive Absorption
CA_TASK	Cognitive Absorption in Task
CA_TECH	Cognitive Absorption in Technology
CMC	Computer-mediated Communication
CR	Composite Reliability
DOTA 2	Defense of the Ancients 2
HMD	Head-mounted Display
IAT	Implicit Association Test
ICT	Information and Communication Technology
ILP	Individual Learning Performance
IS	Information Systems
IVR	Immersive Virtual Reality
KS	Knowledge Sharing
LoL	League of Legends
MOBA	Massive Multiplayer Online Battle Arena
NCCPE	National Co-ordinating Centre for Public Engagement
ODE	Online Disinhibition Effect
PT	Perceived Trust
RQ	Research Question
RRI	Responsible Research and Innovation
SCT	Social Cognitive Theory
SDT	Self-Determination Theory
SEM	Structural Equation Modeling
TB	Toxic Behavior
TPB	Theory of Planned Behavior

ValSD	Value-Sensitive Design
VirtSD	Virtue-Sensitive Design
VR	Virtual Reality

Part A: Introduction

1 Overview

1.1 Opportunities of Gamification for Learning

Gamification is becoming an increasing part of our lives, in both the personal and the work life domains. Games are such a highly motivating experience that, in 2020, about two billion people played video games – a number that is expected to rise to three billion by 2023 (Statista, 2020). Consequently, the inclusion of game elements in non-game contexts, in the form of gamification (Deterding et al., 2011), is being increasingly used in businesses and education. With its inclusion of game elements, gamification is expected to increase motivation and performance in many types of non-game activities. In line with these expectations, analysts predict the gamification market to grow from 9.1 billion US dollars in 2020 to 30.7 billion US dollars in 2025 (MarketsandMarkets, 2020). A large part of this growth is due to educational uses of gamification, of which, in particular, augmented reality (AR) and immersive virtual reality (IVR) will be technologies worth watching (MarketsandMarkets, 2019). Thus, as a popular topic with enormous business relevance, it is essential for information systems (IS) research to study this phenomenon.

Research on gamification has shown that it can indeed positively affect learning outcomes. A recent meta-analysis has indicated that there is a small but positive effect of gamification on learning performance as well as motivation (Sailer & Homner, 2020). For example, gamification can enhance students' engagement in schools (Poondej & Lerdpornkulrat, 2016), increase IS use in organizations (Morschheuser et al., 2015), and even heighten performance in work-related areas such as requirements engineering (Lombriser et al., 2016).

However, some research has found negative or ambiguous effects of gamification on learning performance. For example, although gamification can foster motivation, this does not always carry over to enhance performance (Stansbury & Earnest, 2017). Moreover, only quantitative performance, and neither qualitative performance nor motivation, was enhanced in a study of a crowdworking task using leaderboards, levels, and points (Mekler et al., 2017). Other studies have found no effects on performance in relation to cognitive learning outcomes, but greater engagement with the learning environment and increased qualitative performance (Hew et al., 2016). Therefore, the question remains why these apparently conflicting effects occur.

The contradictory findings about gamification suggest that design and implementation as well as the context of gamification play a crucial role in improving performance and that sufficient theoretical knowledge is still lacking (Seaborn & Fels, 2015). Moreover, it is still unclear which specific gamification elements are effective, how they can be combined, and how the implementation of specific gamification design elements might affect learning-related outcomes. Therefore, the aims of this thesis are to a) identify and test the effects of gamification design elements and b) gain deeper insights into the contexts in which gamification takes place.

1.2 Research Background

This thesis investigates gamification effects in relation to learning. Learning reflects "changes in the behavior of an organism that are the result of regularities in the environment of that organism" (De Houwer et al., 2013, p. 631) and, as such, is an integral part of IS research. IS research has looked at a range of topics in the context of learning, including the improvement of individual learning (Bostrom et al., 1990), facilitation of organizational learning (Alavi, 1994), and design of effective multimedia learning (Coppola et al., 2002).

The effects of gamification on learning can be viewed from different theoretical perspectives. First, the traditional perspective of technology acceptance, in the form of the technology acceptance model (TAM, Venkatesh & Davis, 2000; Venkatesh et al., 2003), has been extended by including indicators for motivation, such as enjoyment and cognitive absorption, to facilitate investigating more hedonic IS (Lowry et al., 2013; van der Heijden, 2004). Moreover, recent frameworks in IS (Liu et al., 2017) have called for the inclusion in gamification research of theories that can not only explain reuse intention but also the formation of motivation, such as self-determination theory (Deci & Ryan, 1985; Ryan & Deci, 2020). Although self-determination theory has been applied in relation to gamification (van Roy & Zaman, 2019; Xi & Hamari, 2019), few experiments have manipulated individual gamification elements to explain the three psychological needs of self-determination theory, with most papers either focusing on only a subset of needs (Mekler et al., 2017) or manipulating more than one gamification element in one condition (Sailer et al., 2017).

To investigate the role of gamification elements that are embodied, such as avatars (virtual bodies), one stream of research in the context of human–computer interaction has focused on the "uncanny valley" effect (Kätsyri et al., 2015; Mori, 1970). This effect states that entities are perceived less favorably when they reach a high but imperfect level of human likeness. In contrast, the "computers are social actors" theory (Nass et al., 1994) states that increased human likeness could facilitate social interaction because humans treat digital humans similar to other humans, regardless of their actual origin. Therefore, it is still unclear which of these theories apply under which circumstances.

In addition to theories focused on user motivation, this thesis takes a dual-process approach to investigate the antecedents and outcomes of learning processes. According to dual-process theories (Soror et al., 2015; Strack & Deutsch, 2004), human behavior is guided by two systems: the automatic and the reflective system. Whereas the automatic system (also called system 1, impulsive system, or intuitive system) is fast, needs only a few cognitive resources, and reflects automatic associations in the brain that are hard to change, the reflective system (also called system 2) is slow, needs a higher amount of cognitive resources, and can adapt to change easily. Because they can explain human thoughts and behaviors both in general and in relation to learning, dual-process theories are well suited to explain the effects of gamification design elements in relation to motivational and performance-related outcomes of learning. Against the background of the theories described above, this thesis attempts to gain deeper insights into explanations of gamification effects in terms of motivational and performance-related outcomes for four main research tracks: 1) gamification design elements, 2) virtual bodies, 3) gamification contexts, and 4) ethical implications.

2 Thesis Structure

This thesis is organized into three parts. Part A gives an overview of the topic, introduces the main theories used, and summarizes the questions and results of 14 research papers. Part B comprises the 14 papers (see Table 2.1 for an overview of the research papers). Part C concludes. A summary of the parts and sections can be seen in Figure 2.1.

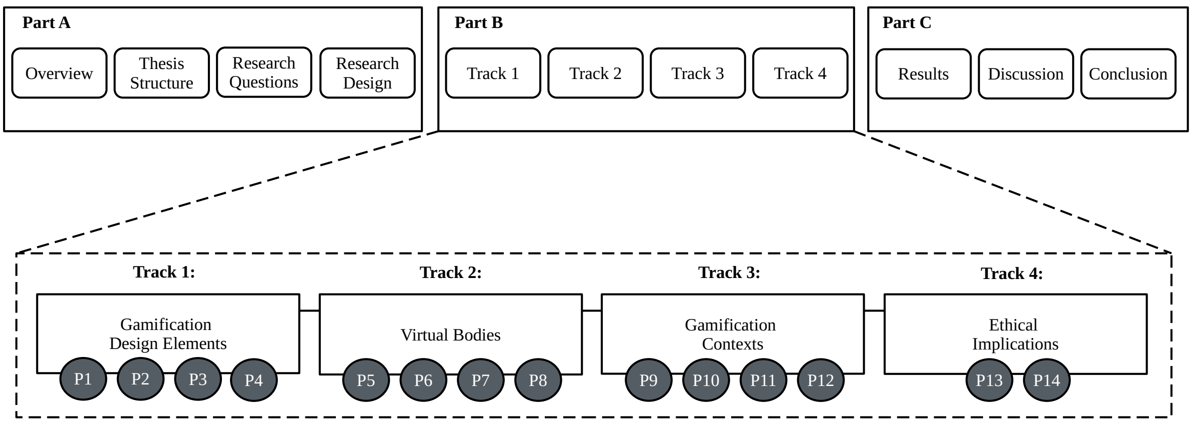


Figure 2.1: Overview of the Thesis

ID	Title	Impact Factor/VHB-JOURQUAL3
Track 1: Gamification design elements		
P1	Jahn, K., Kordyaka, B., Machulska, A., Eiler, T. J., Gruenewald, A., Klucken, T., Brueck, R., Gethmann, C. F., & Niehaves, B. (2021). Individualized gamification elements: The impact of avatar and feedback design on reuse intention. *Computers in Human Behavior*, *119*, 106702	6.829
P2	Jahn, K., Oschinsky, F., Kordyaka, B., Machulska, A., Eiler, T. J., Grünewald, A., Brück, R., Klucken, T., Gethmann, C. F., & Niehaves, B. (in revision). Design elements in immersive virtual reality: The impact of object presence on health-related outcomes. *Internet Research*	6.773
P3	Jahn, K. (2019). Gamification elements in immersive virtual reality. Comparing the effectiveness of leaderboards and copresence for motivation. *Proceedings of the 9th International Conference on Advanced Collaborative Networks, Systems and Applications (COLLA),* 8–11*, Rome, Italy*	–
P4	Jahn, K., Kordyaka, B., Scholz, T., & Niehaves, B. (2020). Gamified helping? The impact of individualized and group-level cooperative evaluation on knowledge sharing. *Proceedings of the 15th Wirtschaftsinformatik 2020 (WI2020), Potsdam, Germany*	C
Track 2: Virtual bodies		
P5	Nissen, A., & Jahn, K. (2021). Between anthropomorphism, trust, and the uncanny valley: A dual-processing perspective on perceived trustworthiness and its mediating effects on use intentions of social robots. *Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS-54), Maui, Hawaii, USA*, 360–369	C
P6	Jahn, K., & Nissen, A. (2020). Towards dual processing of social robots: Differences in the automatic and reflective system. *Proceedings of the 41st International Conference on Information Systems (Virtual ICIS), Hyderabad, India*	A
P7	Jahn, K., & Kordyaka, B. (2019). The effects of robotic embodiment on intergroup bias: An experiment in immersive virtual reality. *Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm, Sweden.*	B
P8	Jahn, K., Kordyaka, B., Ressing, C., Roeding, K., & Niehaves, B. (2019). Designing self-presence in immersive virtual reality to improve cognitive performance - A research proposal. *Proceedings of the NeuroIS Retreat 2019, Vienna, Austria*	–

Track 3: Gamification contexts		
P9	Kordyaka, B., Jahn, K., & Niehaves, B. (2020). Towards a unified theory of toxic behavior in video games. *Internet Research*, *30*(4), 1081–1102	6.773
P10	Jahn, K., Klesel, M., Lemmer, K., Weigel, A., & Niehaves, B. (2016). Individual boundary management: An empirical investigation on technology-related tactics. *Proceedings of the 20th Pacific Asia Conference on Information Systems (PACIS 2016), Chiayi, Taiwan*	C
P11	Jahn, K., Kampling, H., Klein, H.-C., Kuru, Y., & Niehaves, B. (2018). Towards an explanatory design theory for context-dependent learning in immersive virtual reality. *Proceedings of the 22th Pacific Asia Conference on Information Systems (PACIS 2018), Yokohama, Japan*	C
P12	Jahn, K., Heger, O., Kampling, H., Stanik, K., & Niehaves, B. (2017). Designing for knowledge-based familiarity, trust, and acceptance: The case of affective technology. *Proceedings of the 25th European Conference on Information Systems (ECIS 2017), Guimarães, Portugal*	C
Track 4: Ethical implications		
P13	Heger, O., Jahn, K., Mueller, M., & Niehaves, B. (2017). Making use of facebook comments for upstream engagement: A systematic approach. *CEPE/ETHICOMP 2017. Turin, Italy*	–
P14	Jahn, K., Kempt, H., Eiler, T. J., Heger, O., Gruenewald, A., Mach, A., Klucken, T., Gethmann, C. F., Brueck, R., & Niehaves, B. (2020). More than ticking off a checklist? Towards an approach for quantifying the effectiveness of responsible innovation in the design process. *Workshop Ethik und Moral in der Wirtschaftsinformatik 2020, Potsdam, Germany*	–

Table 2.1: The Research Papers Used in the Thesis

2.1 Part A

Part A introduces the gamification literature and provides an overview of the topics and results of the research papers used in this thesis. The introduction above briefly described the motivation for the thesis topic (section 1.1) as well as definitions and how gamification relates to IS research (section 1.2). The present section (section 2) describes the thesis structure. The remainder of part A introduces the research questions (section 3) and methodological approaches (section 4) of the four main research tracks (gamification design elements, virtual bodies, gamification contexts, and ethical implications).

2.2 Part B

Part B consists of fourteen research papers (two published journal articles: eleven published conference articles and one journal article undergoing revision). This part is differentiated into four main research tracks, which categorize the papers into coherent sections. Track 1 describes the effects of gamification design elements on motivational and performance-related outcomes. Track 2 focuses on a specific gamification design in the form of virtual bodies and investigates the effects on learning-related outcomes. Track 3 gives more insights into the contexts in which gamification can take place in both the personal life and work life domains. Finally, track 4 identifies ethical implications in relation to gamification and related contexts. An overview of the presented papers can be seen in Table 2.1.

2.3 Part C

Part C concludes this thesis with a discussion of the papers described in part B. Specifically, section 9 summarizes the results of the papers in part B. The results are then discussed in relation to theoretical (section 10.1) and practical implications (section 10.2) as well as possible limitations and ideas for future research (section 10.3). Finally, part C closes with a conclusion of the added value for gamification research and learning theories for individualized learning systems (section 11).

3 Research Questions

This section gives an overview of the research questions associated with the four main research tracks of the thesis.

3.1 Track 1: Gamification Design Elements

The first research track aimed to identify gamification design elements that facilitate learning-related outcomes. Learning-related outcomes can be differentiated into motivational and instrumental outcomes. Whereas motivational outcomes refer to constructs related to short-term or long-term engagement with a learning system, performance-related outcomes refer to whether the actual performance improves after using the learning system. Previous research on gamification has highlighted that both outcome types should be considered in research on gamification (Liu et al., 2017).

Although there is an overall positive effect of gamification on motivational and performance-related outcomes (Sailer & Homner, 2020), studies have also found mixed results. Possible explanations for these mixed results are a) reduced comparability of effects because of varying contexts, b) methodological issues, and c) lack of knowledge about the theories (Seaborn & Fels, 2015). To address this problem, track 1 draws on theoretical knowledge to identify and test how design elements affect learning-related outcomes.

Research Question 1. How should gamification elements be designed to increase motivational and performance-related outcomes?

Related to the design of gamification elements and the identification of theoretically sound gamification design elements is the question why specific gamification design elements achieve specific positive or, possibly, negative effects. This lies at the core of design science research and, specifically, explanatory design theories (Niehaves & Ortbach, 2016). Against the background of this methodological approach, track 1 used and tested three theoretical approaches to gain insights into the working mechanisms of gamification design elements. First, self-determination theory (Deci & Ryan, 1985) was used to explain the effects of motivational outcomes against the background of the three psychological needs of competence, social relatedness, and autonomy. Second, dual-process theories (Hofmann et al., 2008; Soror et al., 2015; Strack & Deutsch, 2004) were used to explain how automatic and reflective processes affect motivational and performance-related outcomes in relation to the design of gamification elements. Third, these theories were connected with theories of presence (Reiner & Hecht, 2009; Schultze, 2010; Stevens & Jerrams-Smith, 2001) to explain the specifics of gamification design elements in an IVR environment. Using these theories, track 1 aimed to explain the working mechanisms of gamification design elements on motivational and instrumental outcomes.

Research Question 2. How can the effects of gamification design elements on motivational and performance-related outcomes be explained?

3.2 Track 2: Virtual Bodies

One relevant component of gamification design elements is the design of virtual bodies. Virtual bodies play an important role in serious or hedonic games, for example as avatars that are used by real humans to represent themselves or as virtual bodies that represent embodied digital agents. Virtual bodies in the form of avatars provide a form of identity expression (Taylor, 2002), and using an avatar in the digital space influences behaviors and attitudes in both the digital and real world (Schultze, 2014), opening up a range of possibilities to change human behavior in a desired way. For example, users who embody a virtual body with greater attractiveness display higher self-disclosure than when they embody a virtual body with less attractiveness (Yee et al., 2009). Likewise, participants using avatars with greater height compared to other participants display more confident behavior in the "ultimatum game" (Yee et al., 2009). In contrast, individuals using an avatar or virtual body parts that differ greatly from their real selves identify less with their virtual representation, which can reduce experiential outcomes in game play (Birk et al., 2016; Schwind et al., 2017; Soutter & Hitchens, 2016; Suh et al., 2011). Specifically, perceiving lower similarity between the self and the avatar decreases attachment to the avatar, task diagnosticity (Suh et al., 2011), and intrinsic motivation (Birk et al., 2016; Soutter & Hitchens, 2016). These results pose the question of how similar to the user virtual bodies should be designed to facilitate learning.

Research Question 3. What role does similarity play in motivational and performance-related outcomes?

As argued above, previous research has shown that similarity to the user can play an important role in increasing acceptance (Suh et al., 2011). However, when future learners or users of a specific system are humans, increased similarity may result in increased human likeness, which may lead to the uncanny valley effect (Mori, 1970). The uncanny valley effect states that entities with a high but not perfect degree of human likeness on an emotional, cognitive, visual, or auditory level decreases acceptance of these entities compared with lower or almost perfect human likeness. However, other theories propose that increased human likeness does not reduce acceptance, because humans treat digital humans in a similar way to real humans (Nass et al., 1994). Therefore, the question remains which theoretical approach is more adequate in which circumstances.

Recent theories have suggested that one relevant factor in the uncanny valley effect might be automatic and reflective processes. In line with dual-process theory (Soror et al., 2015; Strack & Deutsch, 2004), recent accounts have identified that both automatic and reflective processes can lead to anthropomorphism (Złotowski et al., 2018). However, there is still little research into how dual-systems theory is related to the acceptance of virtual bodies in an automatic and reflective system. Therefore, an additional aim of this thesis is to gain insights into the processes involved in the acceptance of virtual bodies.

Research Question 4. What role do reflective and automatic processes in anthropomorphism play in the acceptance of virtual bodies?

3.3 Track 3: Gamification Contexts

As previously noted, the contexts in which gamification takes place likely play a major role in explaining the inconsistent findings of gamification effects (Seaborn & Fels, 2015). For this reason, it is useful to assess other design elements and working mechanisms that could be used in combination with gamification design elements. This becomes especially important when looking at emerging technologies such as IVR and affective technology. Therefore, track 3 selected two IVR contexts, one affective technology context, and two traditional settings (video games and smartphones) to gain better understanding of the contexts and possible learning-facilitating design features of contexts in which gamification may take place. Additionally, contexts were selected to reflect both work and personal life domains. Thus, track 3 aimed to identify working mechanisms for learning in different contexts of individualized learning systems.

Research Question 5. How can learning in different contexts of individualized learning systems be explained?

3.4 Track 4: Ethical Implications

As argued in the previous sections, the increasing diffusion of new technologies such as IVR, AR, and affective technology into humans' everyday lives comes with a range of expectations for increasing learning for maximizing health, cognitive performance, and many more areas. However, with this increased complexity of technologies, the design of technologies increasingly requires the consideration of ethical implications. In line with this need, responsible research and innovation (RRI) has been proposed as a transparent process to facilitate the design of socially desirable and acceptable technologies (Owen et al., 2013). One major component of RRI is deliberate activities for involving different stakeholders who can bring a diverse range of perspectives into the design process.

However, particularly for technologies that are not used or known by many people, motivating stakeholders to share their perspectives to identify relevant ethical topics can be challenging. Therefore, track 4 looked at how the ethical implications of relatively unknown technologies can be investigated.

Research Question 6. How can ethical implications for unfamiliar technologies be investigated?

In addition to the involvement of stakeholders, a major problem for designing technologies is that not all design features that are socially desirable in the RRI process are addressed with a specific design. Reasons for this may be failure to identify the design feature, limited resources, or a conscious choice not to implement the feature. Moreover, different RRI approaches exist. For example, whereas value-sensitive design (Friedman et al., 2017; van den Hoven, 2013) focuses on solving value conflicts by involving stakeholders, virtue-sensitive design (Vallor, 2016) highlights the role of educating researchers.

Which of the RRI approaches is the most effective remains an unsolved problem. Moreover, the reasons why an ethical design feature is (or is not) implemented may be unclear if researchers report only the identified norms or values and the final design artifact. This makes it especially difficult to identify whether a method aimed at achieving a more ethically sound design leads to better, equal, or worse results than another method. Therefore, the second research question in track 4 aimed to find solutions to increase transparency in the design processes and outcomes.

Research Question 7. How can design choices with ethical implications be made transparent in the design process?

4 Research Design

This thesis uses various methodological approaches to investigate the research questions developed above. First, the research papers can be differentiated into full and short papers on conferences and full papers published in or submitted to journals. The short papers focus on reviewing existing theories and empirical results in combination with a research proposal with the aim of stimulating future research. They also provide an excellent way to discuss planned research with other experienced researchers and therefore include their comments prior to conducting the actual research, which can increase the quality of planned studies. In this thesis, the different tracks start with one or multiple full papers that describe a completed study and are further complemented by short papers functioning as a basis for the completed studies or suggesting future research by making a theoretical contribution. Notably, when a short paper serves as a basis for a full paper, the full paper may deviate from the short paper due to feedback from presentation of the paper at the respective conference. However, short papers give insights into not-yet conducted research at the end of a track, to highlight specific relevant results of the literature review or relevant research questions for further investigation of the topic.

Regarding the methods used, the papers can be differentiated in two major approaches: 1) quantitative, in the form of experiments and surveys, and 2) qualitative research, in the form of case studies. Quantitative studies in the form of surveys were analyzed using regression analysis and path analysis. For experiments, the analysis methods were more diverse and included analysis of variance (ANOVA) and analysis of covariance (ANCOVA), both in repeated measures and in non-repeated measures variants. Additionally, more complex analyses included multilevel mediation analysis and between-subjects experimental data analyzed with structural equation modeling techniques. Furthermore, the measurement of dependent variables differs. Whereas some studies used only questionnaire data, others combined questionnaire data with actual learning performance measured through behavioral information, a learning test, or reaction-time information, each providing a more objective source of information than purely subjective questionnaire data. Finally, the method of participant recruitment and treatment of participants differed: Whereas most studies used a convenience sample in a laboratory or online setting, some studies also used more representative samples via recruitment on crowdworking platforms (e.g., clickworker).

The qualitative studies used different collection and analysis methods. First, grounded theory based on interviews in a case study was used. Second, sentiment detection based on online social media platform comments was conducted to develop a new method for the early engagement steps of RRI. Third, a case study based on ethics workshops was applied to develop a method that increases the transparency of design choices to compare the effectiveness of various RRI approaches.

Part B: Research Articles

5 Track 1: Gamification Design Elements

5.1 Paper 1: Individualized Gamification Elements

Title	**Individualized Gamification Elements: The Impact of Avatar and Feedback Design on Reuse Intention**
Authors	Katharina Jahn katharina.jahn@uni-siegen.de Bastian Kordyaka bastian.kordyaka@uni-luebeck.de Alla Machulska alla.machulska@uni-siegen.de Tanja Joan Eiler tanja.eiler@uni-siegen.de Armin Gruenewald armin.gruenewald@uni-siegen.de Tim Klucken tim.klucken@uni-siegen.de Rainer Brueck rainer.brueck@uni-siegen.de Carl Friedrich Gethmann carl.gethmann@uni-siegen.de Bjoern Niehaves niehaves.bjoern@uni-siegen.de
Publication Type	Journal Publication
Publication Outlet	Computers in Human Behavior
Outlet Information	Impact Factor: 6.829
Status	published
Full Citation	Jahn, K., Kordyaka, B., Machulska, A., Eiler, T. J., Gruenewald, A., Klucken, T., Brueck, R., Gethmann, C. F., & Niehaves, B. (2021). Individualized gamification elements: The impact of avatar and feedback design on reuse intention. *Computers in Human Behavior*, *119*, 106702

Table 5.1: Fact Sheet Paper 1

Individualized Gamification Elements: The Impact of Avatar and Feedback Design on Reuse Intention

Abstract

Gamification is often equipped with the promise to increase motivation and performance. However, research explaining which gamification design elements are effective and the mechanisms through which these effects can be explained is still at an early stage. By drawing on the three psychological needs – competence, autonomy, and social relatedness – proposed by self-determination theory, we develop a model to explain the effects of feedback and avatar design on reuse. We test these effects with a 2 (avatar similarity: low vs. high) × 2 (embodied feedback: no feedback vs. embodied feedback) × 2 (status feedback: no feedback vs. score and leaderboard) + 1 (control group) experiment. Additionally, we use structural equation modeling to test the derived model. Our results support evidence that different forms of feedback and avatar design influence reuse by satisfying the three psychological needs. Furthermore, our findings also reveal that autonomy for decision freedom is negatively related to reuse intention, which disagrees with existing research and may provide insights into why results on gamification elements are inconsistent.

Highlights

- Reuse intention was analyzed in an experimental approach using a gamified version of the manikin task
- Gamification elements satisfy reuse intention and the psychological needs through different working mechanisms
- The three psychological needs show unique relationships towards reuse intention
- Differentiating between autonomy of decision freedom and task meaningfulness reveals converse correlational patterns

Keywords— gamification, self-determination theory, embodied feedback, status feedback, avatar similarity, manikin task

5.1.1 Introduction

In recent years, gamification has received increasing attention from scholars and practitioners as a tool for motivating users to increase their health and performance. Gamification describes the incorporation of game design elements into an existing context or system to achieve desired outcomes (Deterding et al., 2011; Liu et al., 2017; Schöbel et al., 2020) and has been applied in a variety of contexts, including learning (Landers & Landers, 2015; Legaki et al., 2020; Shi et al., 2014), work (Arai et al., 2014; Fernandes et al., 2012), ecommerce (Xi & Hamari, 2020), cognitive tasks (Groening & Binnewies, 2019), crowdsourcing (Liu et al., 2011), and health (Lister et al., 2014; Lumsden et al., 2016).

Although gamification comes with the promise to increase motivation and performance, gamified systems cannot always live up to this promise (Hamari, 2013; Hanus & Fox, 2015; Mekler et al., 2017; Mitchell et al., 2017) and effect sizes were only small in a recent meta analysis (Sailer & Homner, 2020). Reasons for this may include low comparability of different contexts, constraints in research methods, and insufficient theoretical knowledge (Seaborn & Fels, 2015). Therefore, it is essential to understand the working mechanisms of gamification elements in order to understand how gamification should be designed to facilitate beneficial outcomes.

Research has shown that different users prefer different gamification elements (Schöbel et al., 2017), depending on the individual characteristics of the user, such as gender (Codish & Ravid, 2017). Different taxonomies of gamification elements also exist (e.g., Robinson & Bellotti, 2013; Schöbel & Janson, 2018), and recent research has identified design principles for gamification elements in relation to user preferences and task congruence (Liu et al., 2017). However, explanations of these effects using underlying theories are still at an early stage.

Self-determination theory (SDT) (Ryan & Deci, 2017; Ryan et al., 2006) is a promising approach to explain the working mechanisms of gamification in relation to motivation and performance. SDT has been applied successfully to propose relationships between gamification elements and the psychological needs of competence, social relatedness, and autonomy in tasks that are highly repetitive (e.g., picture annotation tasks (Mekler et al., 2017)) as well as more complex tasks (e.g., simulation of an order-picking process (Sailer & Homner, 2020)). However, findings are still mixed regarding specific combinations of gamification elements and needs (Mekler et al., 2017; Sailer et al., 2017). Additionally, research has mainly investigated gamification effects related to SDT without attempting to identify relationships between the three needs. Therefore, there is a gap in the need to explain and compare the effectiveness of gamification elements in relation to SDT and reuse. Consequently, our paper is guided by the following two research questions:

Research Question 1. How can different forms of feedback and avatar design influence psychological needs and reuse intention?

Research Question 2. How do the three psychological needs relate to each other in the initial phase of adoption of a gamified task?

To answer our research questions, we conducted a 2 (avatar similarity: low vs. high) × 2 (embodied feedback: no feedback vs. embodied feedback) × 2 (status feedback: no feedback vs. score and leaderboard) + 1 (control group) between-subject experiment, and use structural equation modeling to test the proposed model. Specifically, we use the utilitarian context of a highly generic cognitive bias modification training from the psychology domain. We experimentally investigate the effects of different gamification design elements on the psychological needs and reuse intention as well as their interplay. As independent variables, we use three design elements with different levels of diffusion in existing research. First, we select the similarity of the user and avatar as a design element that has been used successfully in previous research (Suh et al., 2011) and is in line with the personalization principle in gamification research (Liu et al., 2017), but has not yet been investigated in relation to theories focusing on the psychological needs. Second, we select embodied feedback as a design element that relates to the basis of the task congruence principle, but has rarely been used in research on gamification (Liu et al., 2017). Third, to test whether the other two design elements can provide additional explanatory power to already established (Sailer et al., 2017; Zhang et al., 2018) design elements in line with user preferences (Schöbel et al., 2017), we select status feedback in the form of the design elements leaderboard and points, which are widely used in existing information systems research.

Our paper is structured as follows. First, we review the literature and identify hypotheses related to (1) the effects of design elements, (2) the role of the three psychological needs, and (3) reuse intention. Then, we describe the methodological details of our experiment. In the subsequent section, we present our results related to the direct experimental results and the structural equation model. Finally, we discuss our findings in relation to theory and practice, deriving as-yet unanswered questions for future research.

5.1.2 Literature Review and Hypothesis Development

In this section, we describe the selection of the context and develop the hypotheses against the background of existing theories. The research model is shown in Figure 5.1.

5.1.2.1 Self-Determination Theory

SDT (Deci & Ryan, 2000) is one of the most widely used theories of human motivation and personality and has already shown that it can explain hedonic (e.g., enjoyment) and utilitaristic variables (e.g., reuse) (Ryan et al., 2006). Besides other theoretical approaches, SDT captures intrinsic and extrinsic shares of motivation on a continuum. The theory proposes that the basis of human motivation for a specific behavior lies in the fulfillment of three fundamental human needs, namely the need for competence, autonomy, and social relatedness.

1. *Need for competence* describes the degree to which individuals perceive feelings of success in their interaction with the environment (Rigby & Ryan, 2011; Sailer et al., 2017).

2. *Need for autonomy* can be differentiated into autonomy related to psychological freedom and autonomy related to volition (Sailer et al., 2017; Vansteenkiste et al., 2010). Autonomy related to psychological freedom describes the degree to which individuals feel that they can decide which actions they want to perform in line with their own values, whereas autonomy related to volition describes how the individual acts out of their own accord, without external pressure. On this basis, autonomy can be differentiated into (a) autonomy for decision freedom, the degree to which the individual perceives that they have the freedom to choose between courses of action, and (b) autonomy for task meaningfulness, the degree to which the individual perceives these actions in alignment with their own values (Sailer et al., 2017).

3. *Need for social relatedness* describes the degree to which individuals perceive feelings of belonging and connectedness to other individuals (Deci & Ryan, 2000).

Research on the role of SDT for gamification is still at an early stage. Up to now, the few studies that applied the SDT in the context of gamification have shown mixed results for the relations between the psychological needs, gamification elements, and gamification effectiveness (Mekler et al., 2017; Sailer et al., 2017; Xi & Hamari, 2019).

In the following section, we draw upon these and related results to select the gamified task and derive the hypotheses.

5.1.2.2 Selection of Task and Gamification Design Elements

We chose to conduct our study with the manikin task (De Houwer et al., 2001), which is mainly used to assess cognitive biases in the context of approach avoidance behavior (e.g., Krieglmeyer & Deutsch, 2010; Neimeijer et al., 2017) and is similar to other approach avoidance tasks used for retraining cognitive biases in the health context but that do not make use of a manikin (Machulska et al., 2016; Schumacher et al., 2016). In the manikin task, individuals see a manikin above or below a picture and have to approach or avoid the picture based on picture content (e.g., high vs. low calorie food) or picture design (e.g., picture was photographed from side vs. picture was photographed from above). Its advantage for gamification research lies in the opportunity to generalize the effectiveness of the gamification elements to other tasks that require a classification decision on the basis of item characteristics displayed on a computer screen. Therefore, the task can be generalizable not only to classical health-related contexts in which the manikin task is traditionally applied, but also to more repetitive tasks as is common in crowdsourcing and used in relation to SDT (e.g., picture classification tasks (Mekler et al., 2017)) and tasks requiring even less knowledge that have been used in gamification research (e.g. repeatedly setting a slider to the value of 50 (Lichtenberg et al., 2020)). Although such tasks do not provide a high degree of intrinsic motivation in itself, they are still able to elicit a certain degree of intrinsic motivation for users (e.g., serving relaxation purposes or providing feelings of success).

To identify suitable gamification design elements, on the one hand, we drew on research that was able to show the effectiveness of gamification design elements in satisfying psychological needs in the context of learning (Mekler et al., 2017; Sailer et al., 2017). Here, the use of badges, leaderboards, and a performance graph increased the degree of satisfaction of the need for competence and the autonomy of task meaningfulness, whereas social relatedness was increased by providing avatars and teammates. On the other hand, we looked at which design elements users preferred. In this aspect, points, goals, levels, feedback, and leaderboards (Schöbel et al., 2017) were the gamification design elements users preferred the most, whereas virtual goods, avatars, time pressure, and loss aversion were the least preferred. Finally, we took into account frequency of use in the context of cognitive bias modification tasks. Here, digital rewards and feedback were the most used design elements.

On the basis of these results, we decided to choose feedback and avatars as context for the design elements in our study, to represent both design elements that are highly used and preferred as well as design elements that are less preferred and used. To represent highly used feedback variants (which we call *status feedback*), we chose the use of multiple presentation of leaderboards and scores during the task. To represent the context of avatars, we decided against using the freedom to choose the avatar, as done in previous research (Sailer et al., 2017), because in an initial step, we wanted to clarify which effects different types of avatar might have in relation to the user, to address and investigate the personalization principle of gamification design (Liu et al., 2017). Thus, we used *avatar similarity* as the second gamification design element. Finally, we chose to investigate a gamification design element that relates to both previous design elements in choosing a form of *embodied feedback* in which users could see negative effects on the avatar if they made a mistake, which could then be reversed by correctly doing the task. Specifically, the body shape of the avatar changed in relation to the user's performance, thus addressing and investigating the task congruence principle of gamification design (Liu et al., 2017), as the training condition of the manikin task we used was designed to practice avoidance of high calorie food. In Figure 5.2, the different conditions are illustrated.

In the traditional version of the manikin task, a one-colored shape of a manikin is used that is more similar to the social role of a male than a female person (De Houwer et al., 2001). Thus, avatar similarity might be low. This is unfortunate as, according to the self-congruity perspective, avatar similarity increases perceived self-congruity, which in turn influences identification with the avatar because individuals are motivated to maintain a consistent self-concept (Suh et al., 2011). In line with this, previous research has shown that high avatar similarity leads to higher identification with the avatar (Suh et al., 2011). Therefore, individuals who visually see similarities between their avatar and themselves (e.g., regarding gender) will experience a higher degree of identification than individuals who do not see such similarities. We therefore hypothesize the following:

Hypothesis 1. High avatar similarity leads to higher identification with the avatar than low avatar similarity.

Embodied feedback provides information about task performance through altered avatar design. Previous research has highlighted that in addition to the individual influencing their own behavior in online and offline contexts, the avatar influences self-perception and identity performance in both contexts (Schultze, 2014). We therefore propose that the gamification element embodied feedback facilitates the process of understanding the task. If the avatar changes its visual appearance in an unfavorable way when the user makes an error, this can help users to understand their task performance more intensely than traditional ways of feedback (e.g., an error message). We therefore hypothesize that users perceive an avatar that gives embodied feedback about the user's performance to be more helpful in evaluating the task performance than no avatar.

Hypothesis 2. Embodied feedback leads to higher perceived avatar diagnosticity than no embodied feedback.

Previous research has theorized that scores and leaderboards influence the degree of satisfaction of the need for competence by providing feedback (Sailer et al., 2017). We propose that the underlying working mechanism of the relationship between satisfaction of the need for competence and status feedback lies in the characteristic of status feedback, namely, giving users information about their performance compared with other groups or individuals. Therefore, status feedback should increase users' perception that the task provides them with information about their performance. Consequently, status feedback increases the individual's understanding of their task performance while doing the task (task diagnosticity) and therefore addresses the need for competence.

Hypothesis 3. Status feedback in the form of scores and leaderboards leads to higher task diagnosticity than no status feedback.

Additionally, we expect that identification, avatar diagnosticity, and task diagnosticity are related. Avatar diagnosticity describes a characteristic of the avatar that helps the user to fulfill the task. Therefore, the experience of fulfilling the task with the avatar in an instrumental fashion should lead to experiencing positive events because of the avatar's characteristics. Consequently, we expect that these positive experiences of avatar diagnosticity lead to higher identification with the avatar. Avatar diagnosticity should also increase overall task diagnosticity because being informed about task performance by the avatar contributes to the information provided about the task. In this respect, avatar diagnosticity constitutes a more specific variant of task diagnosticity. Consequently, we do not hypothesize that it will directly affect social relatedness or competence because the variance explained by avatar diagnosticity should be due to changes in task diagnosticity or identification.

Hypothesis 4. Avatar diagnosticity is positively related to (a) identification and (b) task diagnosticity.

5.1.2.3 Matching Feedback and Avatar Design to Psychological Needs

Previous research related to SDT in the gamification area has often investigated the combined effect of the needs with almost the same predictors and outcomes for each need and without proposing specific relationships between the needs (Sørebø et al., 2009; Xi & Hamari, 2019). We argue that treating the three needs separately has two advantages, especially in the context of gamification. First, research has shown that specific gamification elements have distinct effects on the three psychological needs (Sailer et al., 2017). By proposing the same relationships between the needs, the opportunity to distinguish between them becomes lost. For example, it is very likely that the social relatedness need is satisfied by means other than the competence need. Second, research investigating the outcomes of the psychological needs have tended to find that the three needs affect outcomes to a different degree, for example in relation to reuse intention (Przybylski et al., 2010; Ryan et al., 2006; Sørebø et al., 2009). However, research in the area of gamification has hardly addressed the question of how the three needs can be distinguished in different phases of information technology adoption and continuance. Therefore, in this paper, we focus on how the psychological needs can be differentiated in the phase of initial adoption (i.e., first use) of the system, and derive specific hypotheses on the relationship of the three needs in this phase below.

Identification with others is a fundamentally relevant aspect for satisfying the need for relatedness. People who identify with entities that look similar or help them in completing a task experience a high degree of emotional attachment to them (Suh et al., 2011). In the context of the manikin task, it is necessary for the avatar to enact the correct movements in line with participants' directives given via keyboard input. On the other hand, the user is responsible for both the avatar's and their own success by giving the correct directions. Thus, using the avatar in the manikin task can be perceived as a form of collaboration with the avatar. Because the social relatedness need can refer to both receiving and providing concern (Deci & Ryan, 2000; Ryan & Deci, 2017), we propose that this experience of collaborating on a task with an avatar that one identifies with can increase the satisfaction of the need for social relatedness (hypothesis H5a).

Additionally, we propose that identifying highly with an avatar in a gamified system also leads to greater experiences of success and failure when using the system. When a user identifies highly with an avatar, making a mistake and having lower scores will be more threatening to their perception of their own competence. Likewise, experiencing success in conducting a task will be more beneficial for the self. This is in line with previous research that showed that higher identification in a vicarious experience leads to higher feelings of self-efficacy (Kang et al., 2021). Therefore, we propose that a higher degree of identification should lead to higher feelings of competence because the experiences of the avatar in the game are more readily transferred to the user when identification is high (hypothesis H5b).

Hypothesis 5. Identification with an avatar is positively related to (a) social relatedness and (b) competence.

In their research on the effect of different gamification elements, Sailer et al. (2017) showed that, among other factors, social relatedness is influenced by the inclusion of a leaderboard. We hypothesize that this relationship can be explained by an increase in task diagnosticity (hypothesis H6a), because a leaderboard provides information about how one's performance relates to other people playing the game. This should be facilitated even further if the leaderboard is presented multiple times, with some names occurring more than once, as is usual in most gaming experiences. Therefore, with the repeated presentation of a leaderboard, individuals can get the feeling of playing with others multiple times, thus creating the experience of competing with other individuals and allowing them to compare how they have performed on similar tasks.

Additionally, with regard to satisfying the need for competence, Sailer et al. (2017) proposed that this can be addressed by providing information on task performance. We therefore hypothesize that higher task diagnosticity leads to increased feelings of competence because performance in the task can be evaluated more accurately (hypothesis H6b).

Hypothesis 6. Task diagnosticity is positively related to (a) social relatedness and (b) competence.

We argue that satisfaction of the needs for autonomy of task meaningfulness and autonomy of decision freedom reflects overall satisfaction of the needs for competence and social relatedness when looking at the early stage of adopting an information system that does not thwart one of the needs to an extreme degree – which is likely the case in many situations in which gamification is employed. As a consequence, autonomy of task meaningfulness and autonomy of decision freedom are influenced by the degree to which the needs for social relatedness and competence are fulfilled. We explain this argument in more detail below.

Regarding the effects of social relatedness, increased perception of this need due to gamification elements (e.g., by seeing names of other people in the leaderboard or by feeling connected to the avatar) allows individuals to experience being connected to others while doing a task. Satisfying the need for social relatedness can then contribute to satisfying the need for autonomy of task meaningfulness. Likewise, social relatedness can increase the need for autonomy of decision freedom by providing an environment with familiar agents (i.e., human users on a leaderboard or an avatar with high similarity) in contrast to a completely digitized environment. We expect that acting in an environment with agents to which individuals feel related leads to higher perceptions of decision freedom because schemata of how to act are more easily accessible. As a consequence, there is a feeling of increased satisfaction of the need for decision freedom.

For the competence need, when individuals feel that they are competent enough to master different situations, they will likely also feel that they can do what they want and that they are able to act in accordance with their own values (i.e., have autonomy). We therefore expect that competence enhances both autonomy of task meaningfulness and autonomy of decision freedom.

Hypothesis 7. Satisfying social relatedness is positively related to autonomy of (a) task meaningfulness and (b) decision freedom.

Hypothesis 8. Satisfying competence is positively related to autonomy of (a) task meaningfulness and (b) decision freedom.

5.1.2.4 Reuse Intention

Although the differentiation between autonomy of decision freedom and autonomy of task meaningfulness has not been applied in many studies, the overall autonomy need has been associated with reuse intention and predictors of reuse in a range of studies (Hsu et al., 2016; Lu et al., 2019; Przybylski et al., 2010; Ryan et al., 2006; Sørebø et al., 2009). In line with this, we also expect to see this relationship in our study. If individuals start to use a new type of gamified system, it is essential for fostering reuse that they perceive a high degree of task meaningfulness. Furthermore, when looking at related theories for explaining reuse intention, such as the technology acceptance model, one of the main predictors for reuse intention is perceived usefulness, which has a weakened relevance for a technology not primarily aimed at utilitarian purposes (e.g., technology aimed at increasing performance at work). In line with this, previous research has found mixed results for the effect of perceived usefulness in these contexts, measured it differently, or found other predictors showing sufficient explanatory power without including perceived usefulness as predictor (Chu & Lu, 2007; Hsu & Lu, 2007; Kim et al., 2007; Liao et al., 2020). As a consequence, Lowry et al. (2013) argue that it is crucial to account for the context when developing perceived usefulness measures. We argue that autonomy of task meaningfulness is well-suited to serve this purpose because it describes the fit of a system with the users' own values – a context well-suited for a gamified system. Likewise, the feeling that one has decision freedom during a gamified task should also enhance reuse intention. Therefore, we hypothesize the following:

Hypothesis 9. Reuse is positively related to satisfying the psychological needs for autonomy of (a) task meaningfulness and (b) decision freedom.

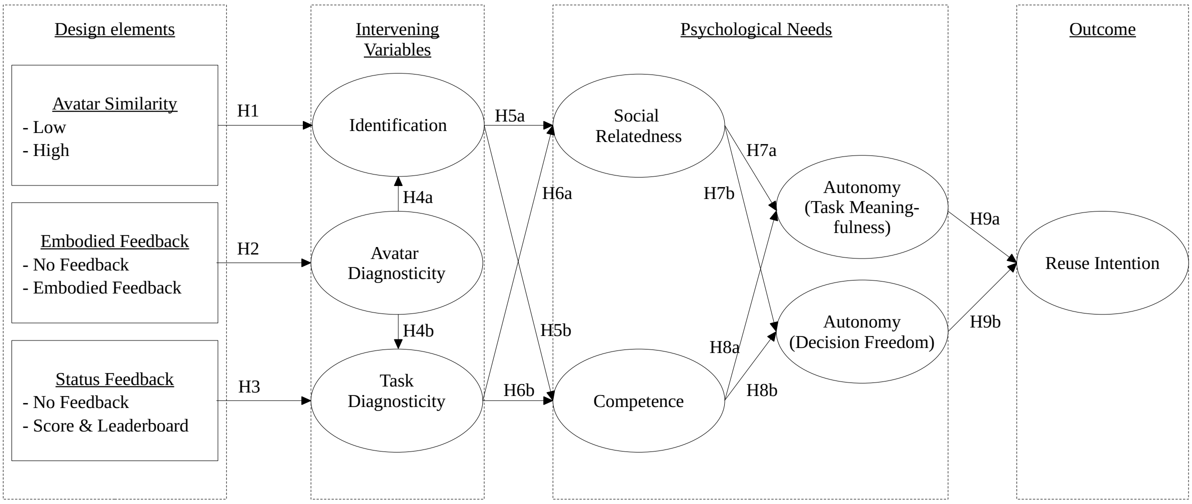


Figure 5.1: Research Model Explaining the Effects of Embodied Feedback, Avatar Similarity, and Status Feedback on Reuse Intention

5.1.3 Method

5.1.3.1 Design and Participants

We used a 2 (avatar similarity: low vs. high) × 2 (embodied feedback: no feedback vs. embodied feedback) × 2 (status feedback: no feedback vs. score and leaderboard) + 1 (control group) between-subjects online experiment to test our hypotheses. We distributed the questionnaire to 342 German-speaking clickworkers[1], who were paid 3 euros. After excluding individuals who completed the questionnaire faster than 97% of

[1]http://www.clickworker.com

the sample, 332 participants remained. Our final sample consisted of 210 males and 122 females (even though participants had the chance to indicate identification with another gender, no one chose to do so). Additionally, 76% of participants reported to have an occupation, and 21% were students. The average age was around 35 years, and participants' mean body mass index (BMI) of about 26 indicated they were slightly overweight. Detailed descriptive statistics for the groups regarding age, BMI, and gender are given in Table 5.2.

		Low Similarity			
		Not Embodied		Embodied	
	Control	No Status	Status	No Status	Status
Age	32.67 (9.31)	37.61 (11.14)	37.82 (13.68)	35.21 (8.80)	35.97 (11.33)
BMI	26.82 (7.78)	24.41 (3.58)	26.31 (5.85)	25.53 (5.05)	24.93 (4.63)
Men	26	18	26	27	21
Women	20	14	8	16	15
		High Similarity			
		Not Embodied		Embodied	
		No Status	Status	No Status	Status
Age		36.50(12.20)	37.03(12.59)	35.53 (12.41)	32.69 (12.73)
BMI		25.06 (4.01)	25.30 (4.68)	24.41 (3.56)	26.28 (5.57)
Men		23	29	20	20
Women		13	7	14	15

Table 5.2: Means and Standard Deviations (in brackets) for the Sociodemographic Variables (Paper 1)

5.1.3.2 Materials

Manikin Task: A modified version of the manikin task (affective Simon task) implemented with PsychoJS[2] on www.pavlovia.org was used to investigate the effect of different gamification elements on reuse intention. The manikin task (De Houwer et al., 2001) is a measurement of automatic approach tendencies, for example for high calorie food (Neimeijer et al., 2017). The task consists of a set of trials in which participants see a picture of a stimulus together with a manikin on the screen. The manikin is displayed above or below the picture, and participants have to make the manikin approach or avoid the picture by pressing the “up” or “down” key according to the color of the frame. If the picture frame was red, participants had to move away from the picture, whereas they had to approach it if the picture frame was blue. Before each trial, a

[2] https://github.com/psychopy/psychojs

fixation cross was presented for 1,000 ms. Because we conducted the experiment online and participants could not refer to the experimenter when they were doing the task, we displayed the text "ERROR" when participants made an error.

The manikin task consisted of four blocks. In the first block, participants could train how the task worked with 10 pictures. In the second block, without further notification, participants received a set of 40 pictures with either high or low calorie food, half of which had to be approached (blue frame) and the other half had to be avoided (red frame). In the third block, all high calorie food pictures had a red frame and all low calorie food pictures a blue frame, so that participants were trained to avoid high calorie food. This type of modification is similar to training in an approach avoidance task, in which pictures are pushed and pulled with a joystick and increase or decrease in size accordingly (e.g., Becker et al., 2015; Machulska et al., 2016; Schumacher et al., 2016; Wiers et al., 2010). Thus, for the bias measurement, both high and low calorie food images had to be approached and avoided, whereas for training, high calorie food images were always avoided and low calorie food images were always approached. In the fourth block, participants had to complete a final 40 trial bias measurements similar to the second block. For each of the bias measurement and training blocks, the manikin position (above or below picture) and the order of pictures were randomized for every participant, and in each of these blocks, all combinations appeared equally often.

Stimuli: For the 10 trial practice, we used neutral images of office items. As stimuli for the bias measurement and training parts, we used five high calorie food pictures (fries, donut, pizza, croissant, chocolate) and five low calorie pictures (melon, apple, cucumber, tomato, grapes), similar to previous research with the manikin task in the eating domain (Becker et al., 2015).

5.1.3.3 Design Elements

Avatar Similarity: Avatar similarity was operationalized by either displaying an avatar with the same gender (high avatar similarity) or a different gender (low avatar similarity). An example image is shown in Figure 5.2.

Embodied Feedback: The gamification design element embodied feedback was operationalized by providing users feedback about their performance through the body shape of their avatar. The participants were instructed that the avatar could increase weight according to their interaction and all participants started with an avatar that had a normal-weight shape (the smallest shape possible). Subsequently, when participants performed the task, the avatar gained weight when they made an error. When par-

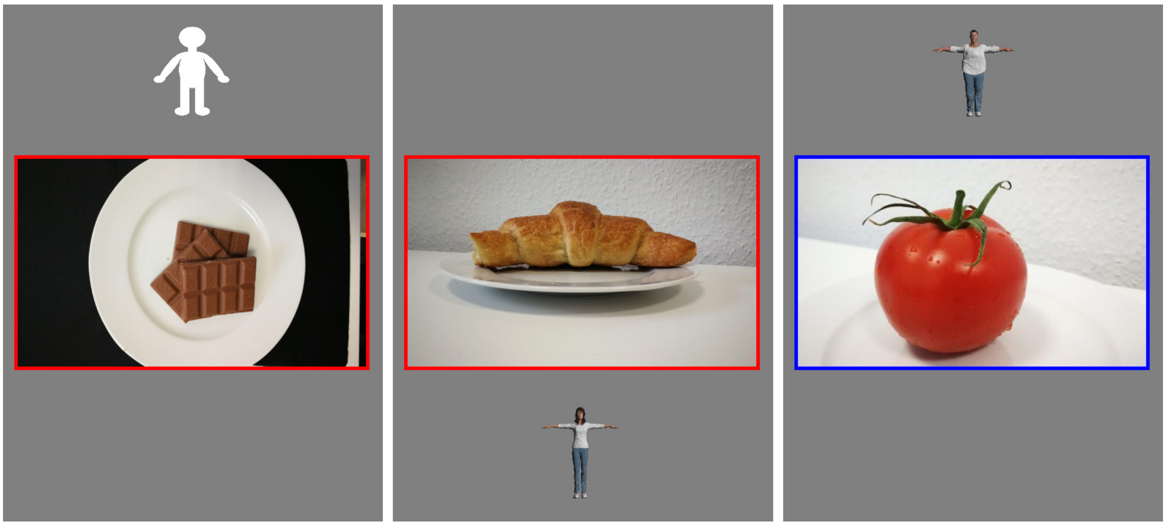


Figure 5.2: Example Images of the Different Conditions (Paper 1). Left: control condition; Middle: normal-weight female avatar; Right: overweight male avatar

ticipants gave a correct answer, the avatar lost weight again. Three body shapes were available (normal weight, slightly overweight, overweight). Therefore, this design is similar to previous research investigating effects of embodied feedback in the health domain (Fox et al., 2009). For participants in the condition without embodied feedback, the avatar always had a normal-weight shape.

Status Feedback: Status feedback was operationalized by displaying the current score and a leaderboard for the pre-bias, treatment, and post-bias measure. Participants in the condition without status feedback received no information in the form of leaderboards or scores on their current performance.

Control Group: The control group was designed similar to the traditional manikin task. Thus, a white manikin was displayed below or above the picture and moved its feet while walking toward or away from the picture.

5.1.3.4 Questionnaire Items

All questionnaire items were measured on a scale from 1 (strongly disagree) to 7 (strongly agree) and adapted from relevant literature. The items are shown in Table 5.7 and items additionally used in further analyses for the study are described in the supplementary material.

5.1.3.5 Procedure

Participants were informed that the study was about the effects of an online gamified task on taste perception of foods. Framing the study in this way had the advantage of avoiding confounding with individual differences in motivation and social desirability compared to framing the study and task with having a specific purpose (e.g., prevention and reduction of overweight). After informed consent was obtained, participants provided information on their gender, age, current job, weight, and height. Next, participants were randomly assigned to a condition and were instructed that the gamified task would start. Additionally, they were informed about the avatar (male, female, or manikin) they were assigned, and completed a manipulation check for avatar similarity.

In the second part of the experiment, the procedure of the gamified task was explained to them. Similar to previous research in the approach bias domain, participants were left unaware about differences in the content of the pictures and received only information about the color of the frame that had to be avoided or approached. Participants first completed a bias measure, then the training, and finally the post-bias measure. In the third part of the experiment, participants completed the questionnaire with the remaining manipulation checks, psychological needs, cognitive absorption, and reuse intention measures. Finally, participants were thanked and debriefed. Participants spent on average two minutes on the first part of the experiment, twelve minutes on the second part, and five minutes on the third part, resulting in an average duration of 19 minutes for the complete questionnaire.

5.1.4 Results

We used analyses of variance (ANOVAs) to test whether our manipulations were successful and to test hypotheses 1–3. Subsequently, we applied structural equation modeling analysis to test hypotheses 4–9. Additional analyses are described in the supplementary materials regarding a) preliminary analyses ruling out possible covariates (cognitive absorption and ease of use) and b) a structural equation model including enjoyment.

5.1.4.1 Experimental Results

The next two sections consecutively describe the manipulation checks of our experiment and the experimental tests of our study. All post-hoc tests were corrected with the Tukey method to account for multiple testing. Additionally, to show the robustness of our hypothesis testing, we compared the effectiveness of the ANOVAs without sociodemographic variables with analyses of covariance (ANCOVAs) in which sociodemographic variables are included as covariates. The means and standard deviations of the experimental conditions are shown in Table 5.3.

		Low Similarity			
		Not Embodied		Embodied	
	Control	No Status	Status	No Status	Status
Identification	3.08 (1.46)	2.31 (1.43)	2.27 (1.30)	2.40 (1.43)	3.47 (1.74)
Dia. Avatar	4.09 (1.47)	2.93 (1.63)	2.89 (1.56)	3.96 (1.75)	4.37 (1.50)
Dia. Task	4.29 (1.58)	4.01 (1.60)	5.25 (1.32)	4.48 (1.41)	5.71 (1.04)
Relatedness	2.58 (1.57)	2.54 (1.43)	2.87 (1.56)	2.52 (1.38)	3.43 (1.72)
Competence	5.03 (1.25)	4.66 (1.12)	5.30 (1.39)	4.94 (1.22)	5.65 (0.93)
Autonomy (Mean.)	4.76 (1.40)	4.51 (1.43)	5.19 (1.28)	4.82 (1.32)	5.40 (1.15)
Autonomy (Deci.)	3.94 (1.97)	3.86 (1.75)	4.80 (1.80)	4.58 (1.73)	4.72 (1.59)
Reuse	4.46 (1.62)	4.09 (1.39)	4.44 (1.43)	4.16 (1.58)	4.94 (1.24)
		High Similarity			
		Not Embodied		Embodied	
		No Status	Status	No Status	Status
Identification		2.91 (1.46)	3.12 (1.57)	3.43 (1.36)	3.01 (1.58)
Dia. Avatar		3.36 (1.46)	2.83 (1.69)	4.24 (1.60)	3.65 (1.66)
Dia. Task		4.28 (1.65)	5.05 (1.57)	4.59 (1.20)	5.40 (1.23)
Relatedness		2.25 (1.35)	2.49 (1.38)	2.97 (1.50)	2.57 (1.50)
Competence		4.88 (1.39)	5.28 (1.05)	5.18 (1.07)	5.34 (1.03)
Autonomy (Mean.)		4.64 (1.20)	5.09 (1.37)	4.80 (1.43)	5.14 (1.20)
Autonomy (Deci.)		4.11 (2.13)	4.93 (1.65)	4.67 (1.53)	4.93 (1.80)
Reuse		4.28 (1.50)	4.55 (1.63)	4.14 (1.50)	4.34 (1.50)

Table 5.3: Means and Standard Deviations (in brackets) for the Manipulation Checks and Dependent Variables (Paper 1)

5.1.4.1.1 Manipulation Checks

Prior to determining our hypotheses, we performed consecutive manipulation checks of our three experimental factors, avatar similarity (high vs. low vs. control), embodied feedback (no feedback vs. embodied feedback vs. control), and status feedback (no feedback vs. score & leadership vs. control), using one-way ANOVAs to test whether the experimental manipulation was successful.

Similarity: A one-way ANOVA comparing the levels of avatar similarity was statistically significant ($F(2, 329) = 73.10, p < .001, \eta^2 = .308$). Post-hoc tests revealed that users in the high avatar similarity condition ($M = 3.88, SE = 0.123$) perceived higher similarity between themselves and the avatar than those in the low similarity condition ($M = 1.79, SE = 0.122, p < .001$) and control condition ($M = 3.06, SE = 0.216, p = .003$). Additionally, the difference between the low similarity and control conditions was significant ($p < .001$).

Embodied Feedback: The conditions varied significantly in a one-way ANOVA for the embodied feedback manipulation check ($F(2, 329) = 91.40, \eta^2 = .357, p < .001$). Post-hoc tests showed that participants in the embodied feedback condition ($M = 4.91, SE = 0.134$) perceived that the avatar was more informational about their performance than participants without embodied feedback ($M = 2.39, SE = 0.139, p < .001$) or in the control condition ($M = 2.75, SE = 0.241, p < .001$). The difference between no embodied feedback and the control was not significant ($p = .410$).

Status Feedback: The one-way ANOVA for the status feedback manipulation check was significant ($F(2, 328) = 101.19, \eta^2 = .382, p < .001$), reflecting that participants in the status feedback condition ($M = 6.09, SE = 0.127$) perceived significantly more that a score was present than those in the no status feedback ($M = 3.66, SE = 0.125, p < .001$) and control ($M = 3.90, SE = 0.224, p < .001$) conditions. The difference between control and no status feedback conditions was not significant ($p = .615$).

All in all, these results indicate that the manipulations of similarity, status feedback, and embodied feedback were successful.

5.1.4.1.2 Hypothesis Testing

To test whether our experiment had the intended consequences, we used the three design elements (avatar similarity, embodied feedback, status feedback) and the three hypothesized predictors (identification with the avatar, avatar diagnosticity, task diagnosticity) in $2 \times 2 \times 2$ ANOVAs and ANCOVAs.

Identification: To test hypothesis 1, we conducted a three-way ANOVA on variable identification. In addition to a significant main effect for similarity ($F(1, 277) = 8.18, p = .005, \eta^2 = .028$), a main effect for the embodied feedback condition ($F(1, 277) = 5.73, p = .017, \eta^2 = .018$), and an unexpected three-way interaction emerged between similarity, embodied feedback, and status feedback ($F(1, 262) = 6.02, p = .015, \eta^2 = .020$). No other significant effects were revealed (all $p > .079$). Including gender, BMI, and age in an ANCOVA did not change this pattern for similarity ($p = .001$), the three-way interaction ($p < .036$), and embodied feedback ($p = .004$), but age showed a significant effect ($F(1, 272) = 11.12, p < .001, \eta^2 = .036$), indicating a positive relationship between age and similarity perception ($b = .026$).

To gain more insights into the interaction effect, we conducted separate analyses for high and low similarity, which revealed that in the high similarity condition, groups did not differ according to embodied or status feedback conditions (all $p > .214$). However, there were main effects in the low similarity condition for embodied feedback ($F(1, 140) = 6.69$, $p = .011, \eta^2 = .041$) and status feedback ($F(1, 140) = 4.29$, $p = .040$, $\eta^2 = .034$), as well as a significant interaction effect between embodied and status feedback ($F(1, 140) = 4.97, p = .027, \eta^2 = .032$). Thus, only the combination of embodied feedback and status feedback ($M = 3.47, SE = 0.248$) led to an increase in identification with the avatar in the low avatar similarity condition compared with the other conditions (see Figure 5.3). Therefore, hypothesis 1 is partially supported by the data.

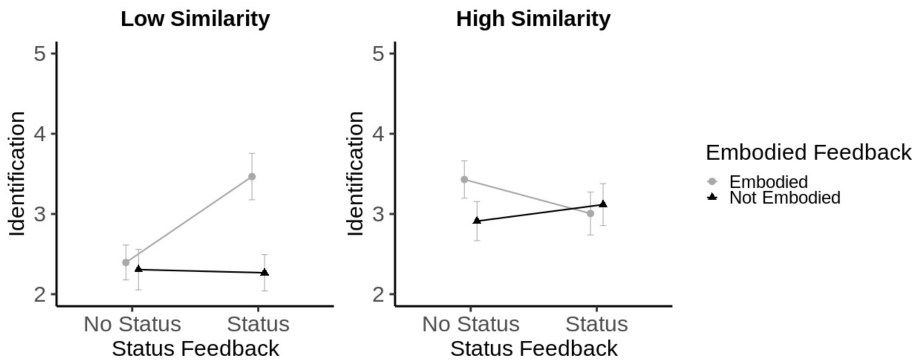


Figure 5.3: Three-way Interaction Effect for Identification (Paper 1)

Avatar Diagnosticity: For hypothesis 2, about the relationship of embodied feedback to avatar diagnosticity, the three-way ANOVA revealed a main effect only for embodied feedback ($F(1, 278) = 30.14, p < .001, \eta^2 = 0.096$, all others $p > .053$). Therefore, embodied feedback ($M = 4.05, SE = .133$) led to higher perceived avatar diagnosticity than no embodied feedback ($M = 3.00, SE = .137, p < .001$). Embodied feedback was still significant when including demographics ($p < .001$), again with a significant positive effect of age ($F(1, 273 = 4.21, p = .041, \eta^2 = .013, b = .017$). All in all, these results support hypothesis 2.

Task Diagnosticity: The ANOVA for task diagnosticity revealed significant main effects for the status feedback condition ($F(1, 278) = 37.68, p < .001$, $\eta^2 = 0.118$) and the embodied feedback condition ($F(1, 278) = 5.80, p = .017$, $\eta^2 = 0.018$), reflecting a higher level of perceived diagnosticity in the status feedback ($M = 5.35$, $SE = 0.117$) and embodied feedback ($M = 5.05$, $SE = 0.117$) conditions than in the conditions without status feedback ($M = 4.34, SE = 0.116$) or embodied feedback ($M = 4.65, SE = 0.119$).

No other main or interaction effects were significant (all $p > .176$), and the inclusion of age, gender, and BMI resulted in only slightly higher effects for the two main effects (status feedback: $p < .001$, embodied feedback: $p = .008$), as well as a significant effect of age ($F(1, 273) = 7.73, p = .004, \eta^2 = 0.026$), reflecting a positive effect ($b = .02$). All in all, these results support hypothesis 3.

Reuse Intention: To assess whether hypothesized effects could still be detected for the last dependent variable, we conducted an additional ANOVA for reuse intention. The ANOVA revealed a significant main effect only for status feedback ($F(1, 277) = 5.21, p = .23, \eta^2 = 0.019$). All other main and interaction effects could not reach significance (all ps $> .195$). Using age, gender and BMI as covariates did not change the significance level of status feedback ($p = .020$) and none of the demographic variables were significant (all ps $> .157$).

5.1.4.2 Structural Equation Modeling

We tested the remaining hypotheses using covariance-based structural equation modeling with maximum likelihood estimation. We used the package lavaan in R (Rosseel, 2012) to conduct the analyses. To assess the validity of our model beyond the comparison of individual fit and inferential parameters, we assessed our main model both with and without control variables, which included the sociodemographic variables age, gender, and BMI in addition to the difference score in performance of the manikin task. The difference score was calculated as the number of correct responses of the second block subtracted from the number of correct responses in the last block, and thus served as a measure of effectiveness of the gamification elements over the course of the task. Including these control variables allowed us to minimize the chance of overlooking possible intervening variables in the structural equation models.

5.1.4.2.1 Measurement Model

The combined investigation of convergent and discriminant validity led to the exclusion of three items (competence1, competence4, reuse3). After exclusion, the standardized lambda values indicated sufficient convergent validity for all lambda and Cronbach's α values (see Table 5.6). Additionally, the average variance extracted (AVE) for each latent construct needs to exceed .50, which was also the case for all constructs (see Table 5.4; note that the square root of the AVE is shown, and therefore all values need to exceed .7).

To assess discriminant validity, we used the Fornell–Larcker criterion (Fornell & Larcker, 1981b), which states that the average variance extracted of a latent variable needs to be higher than the squared correlation between this variable and another variable. This was the case for all constructs (see Table 5.4), with the borderline exception of autonomy for task meaningfulness.

Construct	1	2	3	4	5	6	7	8
1. Identification	**.87**							
2. Dia. Avatar	.70	**.85**						
3. Dia. Task	.39	.40	**.83**					
4. Relatedness	.66	.45	.40	**.86**				
5. Autonomy (Mean.)	.42	.41	.57	.51	**.80**			
6. Autonomy (Dec.)	.24	.11	.28	.25	.42	**.83**		
7. Competence	.40	.34	.47	.46	.79	.37	**.82**	
8. Reuse Intention	.41	.40	.44	.50	.81	.22	.54	**.84**

Table 5.4: Correlations of Constructs in the Solution of the Confirmatory Factor Analysis (Paper 1). The bold numbers display the square root of the average variance extracted.

	Without Control Variables	With Control Variables
χ^2(df)	571.12 (312)	677.67(413)
CFI	.954	.953
RMSEA	.050	.044
SRMR	.055	.052
AIC	20034.28	23647.411

Table 5.5: Fit Indices of the Different Models (Paper 1)

5.1.4.2.2 Structural Model

Model Comparison: The results of the structural model for the psychological needs as separate factors are shown in Figure 5.4. The fit indices are shown in Table 5.5. Although the χ^2-value was significant, the overall model shows a good fit in all indices, regardless of whether or not control variables were included. Additionally, we also included a model with enjoyment in the supplementary material because enjoyment, as one form of intrinsic motivation could be regarded as a mediator. This model shows slightly worse fit compared to our proposed model while complexity is increased.

Path Coefficients: Looking at the path coefficients, the results support all hypotheses except H7b and H9b. Therefore, although the path coefficient for the relationship between social relatedness and autonomy of decision freedom is positive (H7b), the path coefficient was not significant. Additionally, although the path coefficient for autonomy of decision freedom and reuse was significant (H9b), this effect was negative, implying that higher autonomy of decision freedom leads to a decreased reuse intention. We discuss the implications of these results for theory and practice in the next section.

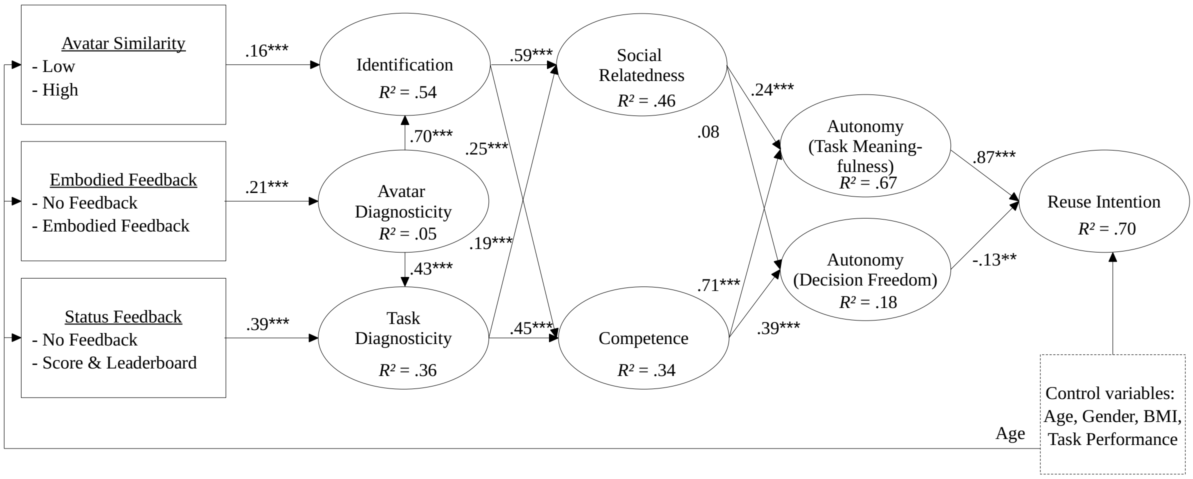


Figure 5.4: Results of the Structural Equation Modeling (Paper 1)

5.1.5 Discussion

5.1.5.1 Implications for Theory

First, our study provides support for the design principles of task congruence and personalization (Liu et al., 2017) and an exemplary case for their application in gamification design. Whereas traditional feedback in the form of scores and leaderboards clearly showed a significant influence on the three psychological needs and reuse intention through task diagnosticity, both avatar similarity (personalization) and embodied feedback (task congruence) were able to provide an independent effect on the dependent variables. Specifically, on the one hand, providing task-congruent embodied feedback influenced the psychological needs and reuse through increased diagnosticity of the avatar. On the other hand, task-congruent embodied feedback and personalization using a similar avatar had an effect by increasing identification with the avatar.

Second, we provide a basis for developing a model to explain reuse and motivation. By explaining the mechanisms through which avatar and feedback design influence the psychological needs and reuse, we have laid the basis for further theorizing to better understand these aspects of gamification. Future research could investigate how additional design elements relate to the ones we have proposed and clarify the role of other related constructs.

Third, we were able to find further support for a positive relationship between psychological needs and reuse and expand the application of SDT in the gamification context by proposing a relationship between the psychological needs. This is especially interesting as we have investigated the need for autonomy in relation to task meaningfulness and decision freedom and proposed that satisfying the need for competence and relatedness influences this meaningfulness perception. As most previous research did not differentiate between task meaningfulness and decision freedom regarding the need for autonomy (e.g., Roca & Gagné, 2008) or did not investigate the relationship between the different needs (Sailer et al., 2017), our results highlight the need to investigate these relationships further in different phases of adoption, both in cross-sectional and longitudinal studies.

Fourth, we were able to differentiate between the action mechanisms that affect how status feedback design elements influence the psychological needs, identified by Sailer et al. (2017). Instead of stating only that specific design elements address the need for competence by giving feedback, with the concepts of avatar and task diagnosticity, we were able to show that this type of feedback influences the psychological needs by not only providing feedback but also by changing users' own perception of how the feedback fits the task. Likewise, by proposing avatar similarity as a design element and identification with the avatar as an intervening variable between similarity and the different needs, we were able to show that providing the user with a specific personalized avatar can have a profound effect on motivation and perception of the self, which is in line with research outside the gamification area (Suh et al., 2011).

Surprisingly, the relationship between autonomy for decision freedom and reuse intention was negative (H9b). Therefore, when individuals perceived that they could decide on their own what to do during the task, they also had less intention to reuse the system. Here, it is important to note that social relatedness did not impact autonomy of decision freedom, whereas competence had a positive impact. Therefore, it might be that autonomy of decision freedom in our study reflects that there is an aspect of competence that, although it leads to more decision freedom, can be inhibitory for reuse intention. For example, it could be that feeling successful in the gamified task leads to a perception of higher decision freedom because of this success. In turn, the task might seem too easy to further satisfy the psychological needs in the future, and therefore leads to decreased reuse intention. This might also explain why we could only find support for a main effect of status feedback on reuse intention, whereas we could find no support for a main effect of avatar similarity and embodied feedback. Additionally, this interpretation could explain why previous gamification research on reuse intention revealed inconsistent results. Conversely, because the p value was high even though it reached significance, the effect could have occurred purely by chance, even though the effect is actually nonexistent or positive in the population. Finally, we have to note that the relationship between autonomy for decision freedom and reuse intention was only negative in structural equation model, whereas it was positive in the correlation matrix. Specifically, there seems to be a unique share of variance in autonomy for decision freedom which is negatively related to reuse intention and is only revealed when removing the common variance with autonomy for task meaningfulness. This result is form of a suppression effect or paradox and may hint towards a mediation or the interference of an unmeasured common cause regarding both autonomy needs. Future research has yet to show what is the case here. On the other hand, this paradoxical effect could also be explained statistically. Here, although we could not find substantial evidence for multicollinearity in the data, it could still be the case that multicollinearity between both autonomy scales or between autonomy for task meaningfulness and reuse intention is responsible for this result. All in all, the unexpected result of the negative relationship between autonomy for decision freedom and reuse intention underlines the relevance of distinguishing between both autonomy needs and highlights that to identify which of these explanations apply, more research is needed, for example with regard to experimental designs that can test specific mediation hypotheses. As a last note, the hypothesis that social relatedness leads to higher autonomy for decision freedom (H7b) could not be supported. Therefore, we cannot be sure whether there is a positive, a negative, or no relationship.

Unexpectedly, our results revealed that avatar diagnosticity was higher in the control than in most other groups. This result might reflect an uncanny valley effect (Kätsyri et al., 2015; Mori, 1970), which describes the effect that a non-human entity might be accepted highly if it slightly human, but when it reaches a certain degree of humanness that is not yet perfect, acceptance suddenly decreases. Therefore, future research could investigate how different features of the avatar's humanness influences avatar diagnosticity, the psychological needs, enjoyment, and reuse intention.

5.1.5.2 Implications for Design

Designers who gamify existing technology can profit from our study in three areas.

First, our study could support the task congruence principle, which highlights the relevance of aligning embodied feedback with the context of the task. Our study was conducted in the context of bias modification for high calorie food, but designers could adapt embodied feedback to other contexts. For example, if a cognitive bias modification task is conducted with the aim of helping individuals to stop smoking, avatars could give feedback on wrong answers by showing typical signs of smoking consequences (e.g., faster aging). This could be especially interesting in the area of virtual reality because of the larger screen sizes and the possibility of creating a high sense of self-presence through body tracking (Schultze, 2010). Second, we were able to show support for the personalization principle with regard to avatar similarity, which might be especially important not only in relation to gender but also to race and other individualized variables. Third, status feedback is the best design element, supported by the fact that it is most frequently used. Nevertheless, other design elements are significant too, and designers could therefore profit from including them.

Finally, any practice has to consider ethical implications of the proposed design elements. Especially in the design of embodied feedback with regard to body shape, promoting a thin body shape could have particularly negative effects on individuals who have eating disorders or issues with physical self-perception. Thus, there is a tradeoff between possible negative effects for some participants and possible positive effects on reuse.

5.1.5.3 Strengths and Limitations

With regard to theory and method, our study has several strengths. First, the experimental design allowed us to detect effects that would have been overlooked in a survey design. Accordingly, we were able to trace back effects on the manipulated levels of the concrete design. Second, we used structural equation modeling, which gave us the chance to capture different theoretical approaches, psychological needs, and design features in one holistic and novel model explaining reuse. Additionally, the use of structural equation modeling allowed us to apply different model tests and fit indicators. This provides a basis for future studies to a) identify additional mediators, b) test mediation effects for every column iteratively and c) unify our model with existing theories for explaining reuse intention (such as technology acceptance models, e.g. (Venkatesh & Davis, 2000)) to further strengthen the explanatory power of our model. Third, the task carried out by participants in our experiment is generalizable to similar tasks (e.g., other areas for cognitive modification bias practices) and classification tasks that could be used for training artificial intelligence, which become increasingly relevant in information systems research.

With regard to our study's theoretical and methodological limitations, we had to make some compromises. First, we collected solely self-reported data, which might have resulted in potential confounds (e.g., social desirability). To increase the validity of our findings, future studies could use behavioral data of participants and compare them with the results of our study. Second, the clickworkers in our sample were paid for doing the experimental task, which may have been a source of potential confounds. Nevertheless, we were able to enhance the needs satisfaction of participants, which strengthened our results. Future studies could ask participants about their motivation to participate in the study and control for the potential of resulting confounds. Third, we need to be cautious regarding the discriminant validity of autonomy for task meaningfulness, external validity of our results and how generalizable the task is to other areas. Accordingly, we recommend investigating neighboring tasks further and looking for similarities and differences. Finally, we manipulated avatar similarity only using gender differences and not body shape information to avoid confounding with the embodied feedback condition, which resulted in rather low perceived avatar similarity and identification ratings. However, if we had changed similarity according to BMI, seeing a change in weight in the first trials would have depended on avatar similarity condition and participants‘ weight. Therefore, future research should investigate the role of avatar similarity and embodied feedback using a non-food context (e.g., smoking) or a higher degree of facial similarity without varying weight.

5.1.6 Conclusion

The current study contributes to literature on gamification and the SDT by providing explanations on the effects and working mechanisms of gamification in relation to technology adoption. Our results showed that both embodied feedback and status feedback influence psychological needs and reuse. Furthermore, we could provide first evidence that autonomy for decision freedom is negatively related to reuse – a result that might explain why previous research on gamification sometimes lead to negative results. Based on our results, future research can investigate whether these relationships remain stable in different contexts, with more elaborate experimental designs focusing on mediational effects, and how the relationships between the three needs might change in different phases of the technology adoption process.

Appendix: Further Information on Measurements

Table 5.6: Factor Loadings of the Different Constructs Used in this Study and Cronbach's α (Paper 1).

Item	Lambdas	Cronbach's α
ident1	.849	.94
ident2	.873	
ident3	.884	
ident4	.908	
ident5	.831	
diaavatar1	.781	.89
diaavatar2	.909	
diaavatar3	.867	
diatask1	.864	.87
diatask2	.783	
diatask3	.843	
sr1	.848	.89
sr2	.870	
sr3	.849	
taskaut1	.820	.84
taskaut2	.827	
taskaut3	.760	
deciaut1	.866	.87
deciaut2	.838	
deciaut3	.803	
competence2	.766	.80
competence3	.877	
reuse1	.903	.82
reuse2	.777	

Table 5.7: Measurement Items (Paper 1)

Manipulation Check: Perceived Similarity (adapted from Fox and Bailenson (2009))
This avatar is similar to me. I think that myself and this avatar resemble each other in appearance. I feel that the avatar resembles me.
Manipulation Check: Status Feedback (self-developed)
I received continuous feedback about my performance through points. My performance influenced my score.
Manipulation Check: Embodied Feedback (self developed)
I received feedback about my performance through the avatar. My performance influenced how the avatar looked.
Need for Competence (adapted from Sailer et al. (2017) and Sailer (2016))
During the gamified task I had feelings of success. I am satisfied with my performance during the gamified task. I felt competent during the gamified task. I felt capable and effective during the gamified task.
Need for Autonomy (decision freedom, adapted from Sailer et al. (2017) and Sailer (2016))
During the gamified task I could decide which actions I execute. During the gamified task I could decide on my own what I do. During the gamified task I could make my own decisions.
Need for Autonomy (task meaningfulness, adapted from Sailer et al. (2017) and Sailer (2016))
It was worthwile doing the gamified task. In retrospect, I feel that it is good to have participated. It was a worthwile experience.

Social Relatedness (adapted from Sailer et al. (2017) and Sailer (2016))
During the task I felt I was part of a team.
During the task I felt socially connected.
During the task I felt emotionally connected to others.
Identification with Avatar (adapted from van der Land et al. (2015))
I had the impression that the avatar was a part of myself.
I identified with my avatar.
I felt connected to my avatar.
My avatar was related to my personal identity.
My avatar appearance was related to my personal identity.
Perceived Diagnosticity of Task (adapted from Jiang and Benbasat (2007))
The task was helpful for me to evaluate my performance.
The task was helpful in familiarizing me with my task performance.
The task was helpful for me to understand my task performance.
Perceived Diagnosticity of Avatar (adapted from Jiang and Benbasat (2007))
This avatar is helpful for me to evaluate my task performance.
This avatar is helpful in familiarizing me with my task performance.
This avatar is helpful for me to understand my task performance.
Reuse Intention (adapted from Lin and Lu (2000))
The gamified task was worthy of doing.
If I could, I would do the task again in the future.
I could imagine to download an app version of the gamified task.

5.2 Paper 2: Design Elements in Immersive Virtual Reality

Title	**Design Elements in Immersive Virtual Reality: The Impact of Object Presence on Health-Related Outcomes**
Authors	Katharina Jahn katharina.jahn@uni-siegen.de Frederike Oschinsky frederike.oschinsky@uni-siegen.de Bastian Kordyaka bastian.kordyaka@uni-luebeck.de Alla Machulska alla.machulska@uni-siegen.de Tanja Joan Eiler tanja.eiler@uni-siegen.de Armin Gruenewald armin.gruenewald@uni-siegen.de Tim Klucken tim.klucken@uni-siegen.de Rainer Brueck rainer.brueck@uni-siegen.de Carl Friedrich Gethmann carl.gethmann@uni-siegen.de Bjoern Niehaves niehaves.bjoern@uni-siegen.de
Publication Type	Journal Publication
Publication Outlet	Internet Research
Outlet Information	Impact Factor: 6.773
Status	in revision
Full Citation	Jahn, K., Oschinsky, F., Kordyaka, B., Machulska, A., Eiler, T. J., Grünewald, A., Brück, R., Klucken, T., Gethmann, C. F., & Niehaves, B. (in revision). Design elements in immersive virtual reality: The impact of object presence on health-related outcomes. *Internet Research*

Table 5.8: Fact Sheet Paper 2

Design Elements in Immersive Virtual Reality: The Impact of Object Presence on Health-Related Outcomes

Abstract

Purpose: Immersive virtual reality (IVR) is frequently proposed as a promising tool for learning. However, researchers commonly implement a plethora of design elements in these IVR systems, making it unclear which specific aspects of the system are necessary to achieve beneficial outcomes. Against this background, we combine the literature on presence with learning theories to propose that the ability of IVR to present 3D objects to users improves the presence of these objects in the virtual environment in comparison with that of such objects as displayed by traditional technologies.

Method: To test our hypotheses, we conducted a 2 (training condition: approach vs. avoid) x 2 (object presence: high vs. low) x 2 (time: pre vs. post) mixed laboratory experiment using IVR with 83 female participants.

Findings: The results support our hypotheses and show that training with high object presence leads to greater reactions to cues during the learning phase (chocolate craving) and improved health behavior (chocolate consumption).

Originality: Our study shows that increased object presence provokes unique experiences for users, which help to reinforce training effects. It sheds further light on how immersive computer technologies can affect a user's attitudes and behavior. Specifically, our work contributes to IVR research by showing that learning effects can be enhanced through an increased degree of object presence.

Keywords— immersive virtual reality, object presence, dual process theories, chocolate consumption

5.2.1 Introduction

The opportunities offered by immersive virtual reality (IVR) have been proposed as a promising way to design effective trainings to promote learning. Especially with the recent availability of IVR technologies for consumers, providing online trainings through IVR has become feasible, and a range of educational IVR trainings already exist for systems such as HTC Vive or Oculus Rift (Ffiske, 2020; Oyelere et al., 2020). The technological characteristics of IVR provide highly realistic scenarios that create the illusion of actually being in the virtual environment (telepresence; Schultze 2010; Slater and Wilbur 1997) while simultaneously offering a high degree of environmental control (Bordnick et al., 2011; Ershow et al., 2011; Hone-Blanchet et al., 2014; Persky, 2011). IVR can thus be used to design enhanced measurements of variables as well as new online training methods that closely resemble real-life exposition scenarios. As a result, IVR can help to improve learning outcomes with respect to areas such as intergroup bias (Bujić et al., 2020), well-being (Osimo et al., 2015), and even performance (Banakou et al., 2018). Additionally, IVR can be used to improve health – an area whose market value will reach 33.72 billion dollars in 2027, according to recent projections (Verified Market Research, 2021).

One major area in which IVR might be especially successful is trainings to reduce and prevent overweightness. More than 1.9 billion adults are overweight, meaning that this major health issue is prevalent among more than one-third of the world population (World Health Organization, 2018). These numbers are alarming because overweightness and obesity constitute risk factors for a range of physiological diseases, including cardiovascular diseases, diabetes, and even cancer as well as mental disorders, such as depression and attention-deficit/hyperactivity disorder (Ahrens et al., 2014). As a result, the direct costs of overweight and obesity are estimated to be approximately 113.9 billion dollars per year in the United States (Tsai et al., 2011). Furthermore, the COVID-19 pandemic may pose additional challenges for healthy eating, especially for individuals who are already overweight (Poelman et al., 2021). These challenges might arise from increased stress during the pandemic, which may facilitate unhealthy eating for already vulnerable groups. Against this background, investigating how the fight against these challenges can benefit from digitization constitutes a salient topic in information systems (IS) research (Agarwal et al., 2010). IS research has already investigated the effects of the digitization of healthcare on experiential and health-related outcomes, such as the doctor-patient relationship (Zhang et al., 2019) and online health communities (Fan & Lederman, 2018; Hur et al., 2019). However, the role of more immersive technologies has only scarcely been investigated in IS, and randomized experiments investigating actual health-related behavior are especially rare.

IVR interventions have been successful in improving outcomes related to food consumption. A major benefit of IVR interventions is that they can more holistically ensure the efficiency and effectiveness of various online interventions in the area of primary prevention (e.g., exercise, nutrition, substance abuse) than similar approaches using desktop computers or smartphones. For example, IVR can contribute to lessening body image disturbances by making cognitive biases conscious through distortions associated with IVR technology (Ferrer-García et al., 2013; Wiederhold et al., 2016). Additionally, by experiencing virtual places and food, users can be exposed to scenarios that are hard to recreate in natural laboratory settings. For example, providing both cues directly related to the task (e.g., an image of a chocolate bar) and cues related to the environment (e.g., a restaurant) in IVR can stimulate a higher level of emotional reaction than that achieved when individuals see only objects on a computer screen (Gorini et al., 2010) and can reduce binge and purge episodes, even in the long term (Ferrer-Garcia et al., 2015; Ferrer-Garcia et al., 2019). These initial findings underscore the great potential of health-related online interventions using immersive computer technologies.

Apart from IVR's rising popularity and relevance, the literature points out that it remains unclear how different immersive computer technologies can affect a population's attitudes and behavior (Bujić et al., 2020). Although previous studies implemented a range of potentially effective design elements, they mainly investigated the overall effect of a complete IVR application in comparison with that of a desktop computer or traditional training (Ferrer-Garcia et al., 2015; Ferrer-Garcia et al., 2019; Gorini et al., 2010). Therefore, it is still puzzling which specific technology-related design elements of IVR are responsible for improving health outcomes and how these effects can be explained theoretically. Specifically, whereas research often proposes the need for a high level of telepresence in a virtual environment that elicits craving, it is unclear what role the virtual design of cues plays in IVR training.

In terms of theoretical approaches to how trainings to reduce overweightness should be designed, research rooted in dual process theories shows that one promising route is the approach avoidance task (AAT). Dual process theories state that human behavior is guided by a reflective system that is deliberate, slow, and flexible and an impulsive system that is automatic, fast, and hard to change. Because health trainings aimed at the reflective system are not effective enough because the impulsive system is hard to change, the AAT is aimed at training the impulsive system. To achieve this, pictures of substance-related cues (e.g., high-calorie food) are repeatedly pushed away on a desktop computer. This type of training has been successfully implemented for reducing consumption of nicotine (Machulska et al., 2016), alcohol (Wiers et al., 2010), and chocolate (Schumacher et al., 2016).

However, even if this research on the AAT on a desktop computer exists, research on design elements of the AAT, especially in immersive settings, is still scarce. Early research in the area of smoking indicates that an AAT IVR training can partly reduce unhealthy behavior relative to the effect of an IVR sham training but does not lead to changes in different measurements of the impulsive system (Machulska, 2021). The inconsistent results of non-immersive and immersive versions of the AAT give rise to the question of which specific design aspects are responsible for these effects. Furthermore, because the study on the IVR version of the AAT (Machulska, 2021) compared only two versions of the AAT training (training vs. sham training) while holding the design of the cues constant, it is still unclear which aspects of IVR design are responsible for positive effects on health-related behavior. Therefore, this paper aims to answer the following research question:

Research Question. What differential impacts does the design of proximal cues in immersive virtual reality have on health-related behavior?

To address this research question, we conducted a laboratory experiment to test the effect of designing cues as 3D food objects (reflecting high object presence) vs. 2D pictures of food (reflecting low object presence). Our results indicate that designing cues with high object presence is sufficient to influence health-related outcomes, even when telepresence remains unchanged. Additionally, the results highlight the relevance of differentiating between processes in the impulsive system when explaining learning in IVR and offer practical insights into developing and applying IVR interventions for the prevention and treatment of overweight.

The paper is structured as follows. In Section 5.2.2, we explain dual process theory and derive explanations for how design elements can be used to influence impulsive and reflective processes using an IVR application that could be distributed as an online training. Building upon this explanation, we use the construct of object presence to explain how the representation of stimuli as 2D or 3D objects affects these processes. In Section 5.2.3, we develop our research model by connecting the theoretical explanations with recent findings on the dual process model. Subsequently, we explain the setup of the experiment in Section 5.2.4 and analyze the results in Section 5.2.5. Finally, we elaborate upon the implications for dual process theory and object presence in the discussion section.

5.2.2 Theoretical Background

5.2.2.1 Dual Process Theory

According to dual process theories (Deutsch & Strack, 2006; Soror et al., 2015; Strack & Deutsch, 2004), human behavior is shaped by the interaction of two different cognitive systems, the impulsive system and the reflective system. The reflective system influences behavior on the basis of deliberately formed intentions. As a consequence, it involves rather slow responses, is goal-oriented and can flexibly adapt to change. In the IS discipline, the reflective system is well known in the form of the theory of planned behavior (Ajzen, 1991) and the theory of reasoned action (Ajzen & Fishbein, 1980), which form the basis of technology acceptance models (e.g., Venkatesh et al., 2003). In contrast to this perspective based on deliberate decisions, the impulsive system is responsible for rapid, automatic responses to stimuli that have been built by associations with successful behavior in the past. As such, the impulsive system is goal-independent and hard to control and change. In IS, the impulsive system has been investigated less extensively but is receiving increasing attention, for example, in relation to decision support systems (Lederman & Johnston, 2011), social networking sites (Polites et al., 2018), and mobile phones (Chen et al., 2019; Soror et al., 2015). Furthermore, a dual process perspective has the potential to explain inconsistent findings in previous research on the effectiveness of virtual reality in learning (Lin et al., 2020). Because the impulsive system is considered hard to change, this paper focuses on the effects of IVR training in this system (see Figure 5.5).

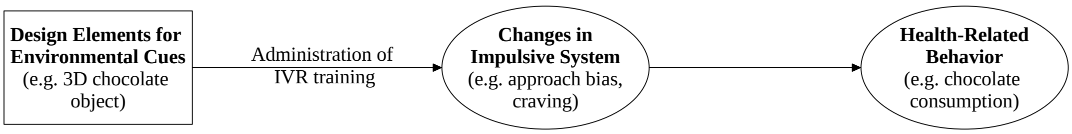


Figure 5.5: Working Mechanism of IVR AAT Training in Relation to the Impulsive System (Paper 2)

One way of explaining how associations are formed in the impulsive system is classical conditioning (van den Akker et al., 2018). From this perspective, when a stimulus is followed by an unconditioned stimulus (US), it can become a conditioned stimulus (CS), which elicits a conditioned response (CR). For example, an individual watching a TV show while eating chocolate may then associate the TV show (CS) with eating chocolate (US). When this individual watches the TV show (CS) at another time, an appetitive response (CR) may be triggered. This process is also called cue reactivity. It is important to note that almost any cue can become an unconditioned stimulus, including the sight, smell, and sound of food as well as internal physiological and psychological states (e.g., hunger, emotions or thoughts) (van den Akker et al., 2018).

When associations in the impulsive system relate to a heightened approach towards certain stimuli (e.g., high-calorie foods), dual process theories propose that individuals show an approach bias. This bias reflects the tendency to physically approach or 'reach out' to these stimuli (Cousijn et al., 2011). As a consequence, psychological measurements based on reaction times to stimuli, such as the AAT, can be used to assess the strength of approach bias. Similar to psychophysiological and neurophysiological NeuroIS measures (Dimoka et al., 2012; Dimoka et al., 2010; vom Brocke et al., 2020), these measures are hard to influence consciously, and thus, they overcome the limitations of self-report measures.

Approach bias towards certain stimuli can be trained by repeated interaction with these stimuli in a desired way (CS-noUS), forming additional associations in the impulsive system. In the context of food consumption, when individuals are trained to repeatedly avoid high-calorie foods, associations to avoid these stimuli are strengthened. As a result, these strengthened associations may weaken existing associations leading individuals to approach these stimuli. Moreover, this form of retraining should be especially effective if training situations are highly similar to consumption situations (van den Akker et al., 2018), for example, with regard to the level of craving that cues elicit. In the context of addiction, a reduction in approach bias after computer-based training using the AAT has been detected for substances such as nicotine (Machulska et al., 2016) and alcohol (Wiers et al., 2013; Wiers et al., 2010). Likewise, in the context of eating high-calorie food, research on chocolate bias has shown that training to push away chocolate images could decrease chocolate bias (Becker et al., 2015; Dickson et al., 2016) or even chocolate consumption (Schumacher et al., 2016). However, some studies could not find a decrease in chocolate consumption (Becker et al., 2015; Dickson et al., 2016), even though they found a decrease in approach bias. On the other hand, some studies showed a decrease in chocolate consumption without a decrease in approach bias (Schumacher et al., 2016). Thus, it is still unclear whether the AAT is actually effective in changing health-related outcomes. Given that these studies were conducted with slightly different operationalizations regarding the content of images, their inconsistent results highlight that it is still unclear what role specific stimuli design elements might play in the AAT. Moreover, recent work has shown that the AAT can be successfully transferred to IVR and partly enhance health-related behavior without affecting several measures of the impulsive system, including approach bias (Machulska, 2021). Against this background, in the next section, we draw upon theories of telepresence and object presence to develop a theoretical understanding of the effects of AAT design in IVR to gain more insights into the role that design elements in IVR training aimed at training the impulsive system can play in health-related outcomes.

5.2.2.2 Telepresence and Object Presence

The sense of being able to act in a virtual space is of great importance to computer-mediated environments (Altschuller & Benbunan-Fich, 2013; Schultze, 2010, 2014; Schultze & Orlikowski, 2010) as well as health-related IVR applications (Tal & Wansink, 2011). Several definitions of telepresence exist, but in a general sense, telepresence can be defined as the "illusion of being in a distant place, that is, being there" (Schultze, 2010, p. 438). The function of the cognitive processes that lead to this illusion is to identify possibilities for acting in the environment (Triberti & Riva, 2016). As such, presence is a subjective experience (Witmer & Singer, 1998) that is influenced by how the technology addresses the users' sensory modalities to recreate reality (Slater & Wilbur, 1997) and cognitive and affective factors (Gorini et al., 2011; Hoffman et al., 1998). For instance, the illusion of 'being there' could refer to the user's sense of actually driving a car by interacting with the throttle or brake, even though they are sitting at a computer in an experimental space.

Acknowledging that virtual and real environments are composed of objects, the definition of telepresence can be rephrased as "the subjective experience of being *co-located with a set of objects*, even when one is physically *not in such a situation*" (Stevens & Jerrams-Smith, 2001, p. 195). Object presence is described as the sense of being able to touch and manipulate a virtual object (Reiner & Hecht, 2009). In line with this, recent definitions of telepresence offer a more agentic perspective by highlighting that presence arises if individuals can engage in behavior based on high-level, conscious intentions in the form of distal future-directed intentions (D-intentions) and proximal present-directed intentions (P-intentions) or automatic motor intentions (M-intentions) in an environment (Pacherie, 2007; Riva, 2009). In relation to dual process theories (Strack & Deutsch, 2004), the reflective system is represented by D- and P-intentions, whereas the impulsive system is represented by M-intentions. Considering these definitions, we define object presence as an individual's subjective experience that a particular object exists in his or her environment that enables him or her to enact behavior on the basis of impulsive or reflective processes. Again considering the illusion of driving a car, a high object presence leads to feeling able to take hold of the (virtual) steering wheel and adjust its height. The object presence in this context is much greater than it would be when presented a photograph in virtual reality that shows a steering wheel – even though neither object actually exists. Although the physical picture will be experienced as a 'real' object, the content of the picture will be perceived as an object only if the individual engages in a high degree of mental simulation. Clearly, although object presence has mainly been researched in actual reality with augmented reality devices (Stevens & Jerrams-Smith, 2001; Sugano et al., 2003), it is becoming an increasingly relevant concept for explaining effects in IVR. Consequently, when creating a task such as the AAT in IVR, the question of how object presence should be addressed arises, especially against a background in which explanations for the effects of IVR on health-related and eating behaviors are still at an early stage (Tal & Wansink, 2011).

5.2.3 Hypothesis Development

Until now, different designs of the AAT have scarcely been considered. Most studies have conducted the AAT using arm movement for initiating approach and avoidance. The arm movements can be implemented by pressing buttons (Roelofs et al., 2009), enacting pull and push movements with a joystick (Chen & Bargh, 1999; Machulska et al., 2016; Wiers et al., 2013) or pushing and pulling a table upward or downward (Foerster et al., 2001). However, arm movements can be ambiguous, as it is unclear whether the stimuli are being pushed away or if the arm is being moved closer to the stimuli. To successfully implement the AAT, it is therefore necessary to have an additional indicator that resolves this ambiguity (Krieglmeyer & Deutsch, 2010). This is usually done by giving visual feedback in congruence with the arm movements. Thus, for push movements, the size of the stimuli decreases, and for pull movements, the size increases. Alternatively, visual feedback can be designed for pull movements in which the distance between the stimuli and a manikin decreases, whereas push movements lead to an increased distance.

From a design perspective, a few new outstanding versions have been developed recently. A version in which the participant is more involved in the task was developed by Stins et al. (2011), who used full-body movements (stepping forward or backward) instead of arm movements to approach or avoid the stimuli. Regarding design elements that vary the presence of the stimulus material, Kim and Lee (2015) developed a virtual reality task for alcohol dependence that showed videos of social situations as stimuli. Additionally, in the area of spider anxious individuals, a conflict version of the AAT has been developed for measurement and training in approaching spiders for spider anxious individuals (Dibbets & Fonteyne, 2015; Mühlberger et al., 2008). Recently, an IVR training with the AAT was developed for smoking (Machulska, 2021). However, an AAT training version in the eating domain that is designed to take advantage of the opportunities offered by IVR is still lacking, although initial evidence suggests that individuals show faster approach reactions to food stimuli than to neutral stimuli (Schroeder et al., 2016). Therefore, we aim to investigate how the design of cues with regard to object presence (i.e., 3D objects vs. 2D images) affects health-related outcomes and compare it with the traditional approach vs. approach condition of the AAT at a desktop computer. The hypothesized effects are explained in the following on the basis of the previously introduced theories.

One reason for the low effectiveness of trainings in the area of substance dependency is low engagement with the training. This is related to treatment participation, which is a prerequisite for positive outcomes (Simpson & Joe, 2004). As it is essential for IS research to investigate both learning performance outcomes and motivation-related outcomes in other domains to avoid overlooking negative effects on acceptance (Liu et al., 2017), we decided to include an indicator of motivation to use the system in the form of enjoyment. Enjoyment is a relevant indicator for assessing engagement with training (Dingle et al., 2008). For telepresence, research has shown quite consistently that telepresence is positively related to enjoyment (Nah et al., 2011; Sylaiou et al., 2010). The reason for this relationship is the higher degree of control and interest that environments with high telepresence provide Nah et al., 2011. We argue that even when telepresence with regard to the environment is held constant, changing merely the presence of the objects the user interacts with results in increased enjoyment because control and interest are still higher for objects that are high in object presence. Specifically, the increased affordances to interact with the objects in the form of D-, P-, and M-intention provide a higher perception of control and interest. Therefore, we hypothesize that object presence can increase the perceived enjoyment of the AAT training.

Hypothesis 1. Individuals in the high object presence condition experience higher enjoyment than individuals in the low object presence condition.

Whereas AAT versions for IVR in the eating domain have yet to be investigated in the area of cue exposure therapy, several versions do exist related to food consumption. In IVR, craving can be increased for different types of cues, e.g., for smoking (García-Rodríguez et al., 2013) or high calorie food (Ferrer-Garcia et al., 2015). One study comparing 3D food objects in IVR with both 2D food images on a computer screen and real food provides evidence that 3D objects in IVR and real food elicit similar emotional reactions in users with anorexia and bulimia nervosa compared with 2D images (Gorini et al., 2010). However, the virtual environment in which the 3D objects in IVR were displayed was a restaurant, which might have confounded the effects of object presence and the effects of contextual stimuli, and neither craving nor food consumption were measured. Additionally, the initial results indicate that displaying 3D objects on a 3D computer screen leads to a comparable level of craving as displaying 3D objects in IVR (Gutiérrez-Maldonado et al., 2016), but in this case, these 3D objects were not compared to 2D images. Therefore, we want to test the effect of object presence alone, as we propose that it is sufficient to elicit a higher level of craving if the stimuli are similar to those encountered by individuals in reality. We thus hypothesize that when object presence is high, food cravings should be higher than when object presence is low.

Hypothesis 2. Individuals who train with stimuli eliciting high object presence experience higher craving than individuals who train with stimuli eliciting low object presence.

Drawing upon classical learning theories (van den Akker et al., 2018), extinction due to cue exposure should be more effective if training situations are similar to real consumption situations. From this perspective, interacting with high object presence stimuli should be more effective than interacting with low object presence stimuli because they provide affordances to interact with objects that are similar to real objects. As a result, 3D objects should be more strongly perceived as objects that could actually be eaten. Because of these affordances, the objects activate behavioral schemata from the impulsive (M-intentions) and reflective systems (D- and P-intentions) more strongly than stimuli in the low object presence condition. In contrast, 2D pictures cannot activate these schemata to the same degree because they do not appear as 3D objects and thus provide the much lower affordances to interact with the object (i.e., although a participant might want to eat the object in a picture, he or she would usually try to grab not the object in the picture but rather the picture itself). In line with this reasoning, previous research has shown that experiencing 3D food objects in comparison with 2D food pictures increases emotional responses to these stimuli after exposure in IVR (Gorini et al., 2010). To illustrate, individuals training with high object presence may have experiences (e.g., high craving) similar to those in situations in which they actually have the opportunity to eat food and behavioral schemata activate. Because they train not to engage in unhealthy behavior (e.g., not to eat high-calorie food) when they have these experiences (e.g., high craving), these training effects are more likely to transfer to a situation in which they can actually eat food and are confronted with these experiences (e.g., high craving) again. On the other hand, when they train with low object presence stimuli for which the training and actual recall situation experiences are not as similar, training effects are less likely to occur. Therefore, it is important to note that H2 and H3 do not contradict each other, even though they may appear to at first sight. We thus hypothesize that repeated interaction with high object presence stimuli without the ability to engage in unhealthy behavior with the object (eating it) acts as a form of extinction by strengthening associations with healthy (non-eating) behavior. As a consequence, unhealthy behavior (e.g., chocolate consumption) should be reduced for users who use the AAT with high object presence stimuli.

Hypothesis 3. Individuals who train with high object presence show less unhealthy behavior than individuals who train with low object presence.

In the investigation of new design elements in the area of health training, it is important to investigate how these elements affect already existing trainings. To achieve this in the case of the AAT, the avoidance training version can be compared to a version in which half or more stimuli are approached instead of avoided. This version should be less effective than the avoidance training version.

Research in the area of actual health-promoting behavior, a result of an interaction of the impulsive and reflective system according to dual process theories Strack and Deutsch, 2004, has shown that health-related behavior can be improved by AAT avoidance training. Studies on cigarettes and alcohol have shown that AAT avoidance training is effective in reducing consumption compared to control training (Machulska et al., 2016; Wiers et al., 2010). Likewise, in the area of food, AAT avoidance training with chocolate images could successfully reduce consumption of, e.g., a chocolate muffin compared to AAT approach training (Schumacher et al., 2016). However, other studies have shown that food consumption in the context of chocolate could not be reduced through AAT avoidance training compared to approach training (Becker et al., 2015; Dickson et al., 2016). These results highlight that an IVR version of the AAT should also investigate how the AAT performs compared to an approach training version.

Regarding the impulsive system in relation to the AAT, the approach bias is an excellent indicator of the impulsive system because it is directly related to the AAT by using the same basic procedures for measurement as the training version of the AAT. Therefore, it can directly measure whether the specific aspects of the impulsive system that have been trained have actually changed. Results regarding approach bias show that in the context of the traditional desktop version, a reduction in chocolate approach bias after AAT avoidance training could be found despite the absence of a significant training effect on alternative snack approach bias after AAT training (Dickson et al., 2016) or in line with a decrease in alternative snack approach bias after alternative snack AAT approach training (Schumacher et al., 2016). Moreover, the results of one study did not indicate changes in the chocolate approach while revealing an increase in stationary object approach bias (Becker et al., 2015).

All in all, these mixed results show the need for additional research in this area. Therefore, we want to test the assumption of the dual process model (Hofmann et al., 2009; Hofmann et al., 2008; Strack & Deutsch, 2004) and hypothesise that training to avoid stimuli with an IVR version of the AAT training should decrease approach bias and health-promoting behavior.

Hypothesis 4. Individuals who train to avoid unhealthy objects show less unhealthy behavior than individuals who train to approach unhealthy objects.

Hypothesis 5. Implicit bias a) decreases after AAT avoidance training and b) increases after AAT approach training.

5.2.4 Method

5.2.4.1 Participants and Design

After recruiting online, over radio, through newspaper articles, and with a physical flyer, 83 participants agreed to participate in the laboratory IVR experiment. In line with previous research in the area of craving and chocolate consumption (Ledoux et al., 2013; Schumacher et al., 2016), we only recruited female participants because previous studies have indicated that women have higher food cravings (Lafay et al., 2001; Weingarten & Elston, 1991) than men and tend to overeat more than men (Burton et al., 2007). Additionally, because previous research has shown that hunger can influence approach bias scores (Seibt et al., 2007), we instructed participants not to drink or eat anything other than water in the two hours prior to the experiment. Finally, they had to like chocolate. Participants were, on average, 27.11 years old (SD = 9.68), and 68.67% were students. Participants on average ate 0.90 chocolate bars per week (SD = .84) and had a BMI of 23.57 (SD = 4.17).

We used a 2 (training condition: approach vs. avoid) x 2 (object presence: high vs. low) x 2 (time: pre vs. post) mixed design in a laboratory experiment. The training condition on object presence varied between subjects, whereas time varied within subjects. To assess the adequacy of the sample size, a power analysis was conducted using G*Power (Faul et al., 2007). Previous research assessing reactions to 3D versus 2D cues in IVR (Gorini et al., 2010) suggested an average effect size of f = .325. Previous research using an IVR version of the AAT in the food domain is not yet available, but developing an IVR application is highly resource intensive. Therefore, we used the medium effect size of f = .325 as an estimator for approach vs. avoidance conditions as well, given that the increased effort required to develop an IVR application would be justified by at least a medium effect size. The results of the power analysis with α = .05 and a power of 80% showed that 80 participants were necessary for a 2 x 2 between-subjects ANOVA. We recruited three additional participants to avoid low power if participants had to be excluded. Groups did not differ significantly in terms of age, chocolate cravings before the AAT training, the number of chocolate bars consumed per week or BMI (all $ps > .15$).

5.2.4.2 Materials

Immersive Virtual Reality. For AAT training, we used an HTC Vive head-mounted display with a resolution of 2160 x 1200, a refresh rate of 90 Hz, and a field of view of 110 degrees. When participants put on the head-mounted display, they saw a room in which they could push and pull either objects (high object presence condition) or pictures (low object presence condition) using the Vive controllers. We used robot hands as the embodiment of the hands, as previous research has shown that they have high presence and likeability and low eeriness ratings (Schwind et al., 2017).

Approach bias measurement. Because the correct measurement of approach bias is highly dependent on correct reaction-time measurements, it was done on a desktop computer with a joystick, a procedure that has been successfully applied in previous research (Machulska et al., 2016). In this measurement AAT version, participants had to repeatedly pull or push pictures with a joystick. If the picture was rotated to the right, participants had to push it away, and if the picture was rotated to the left, they had to pull it. As the picture was pulled, it became larger, and as it was pushed, it became smaller.

The measurement AAT version consisted of three parts: a practice version aimed at increasing participants' familiarity with the AAT, a pre-measure that took place before the IVR training and a post-measure that took place after the IVR training. The pre- and post-measures were necessary because differences between participants' approach bias scores can be considered by calculating difference scores. For the practice version, participants were presented with 10 neutral images that were rotated to the left or right for the pre-measure (2 practice images were presented for the post-measure). For the pre- and post-measures of the AAT, participants had to push and pull a set of 10 chocolate and 10 fruit images. Each picture was presented 4 times in total for each version so that participants pulled and pushed images 80 times for one bias measurement. Participants were instructed to complete the task as fast as possible while making as few errors as possible. Approach bias scores are then calculated by subtracting the reaction time for pulling an image of a specific category (e.g., chocolate images) from the reaction time for pushing an image of the same category. Therefore, higher scores indicate higher approach bias for that category.

Health-Promoting behavior. To measure health-promoting behavior, we measured chocolate consumption during a chocolate tasting. Participants were provided with two bowls of 100 g of chocolate (1 bowl of white chocolate and 1 bowl of whole milk chocolate) and given ten minutes for chocolate tasting. After the tasting, both flavors of the remaining chocolate were weighed. The weight was then subtracted from 100 g to obtain the amount of chocolate that was eaten.

Questionnaires measurements. The scales used in this study are displayed in Table 5.9. Chocolate craving was measured before and after the AAT was conducted using the chocolate experience questionnaire (CEQ). For the manipulation check for object presence, two self-developed items were used. Additionally, a manipulation check for telepresence was used to ensure that only object presence and not telepresence was manipulated.

5.2.4.3 Design Elements

Object presence. The design elements of the cues we presented to participants are displayed in Figure 6.5. Participants in the high object presence condition interacted with stimuli that were similar to real objects while standing in front of a table on which the objects were placed. Participants had to throw objects with a red border into a trash bin that was placed behind the table (avoid condition) and had to put objects with a blue border into a box that was standing directly in front of them (approach condition).

Participants in the low object presence condition stood a few meters away from a striped wall. The stimuli appeared as pictures presented in front of the participant, and they had to pull pictures that were rotated to the left and push pictures that were rotated to the right.

The basic interaction of the AAT task worked identically in both conditions. New stimuli appeared when participants pressed the track pad with their thumb, and stimuli could be moved by moving the controllers towards the stimuli and pressing the trigger to grab them.

Training condition. Participants in the approach condition had to push all fruit stimuli (objects or pictures, depending on the object presence condition) away, whereas they pulled all chocolate stimuli towards themselves. In the avoid condition, they had to push away all chocolate stimuli and pull all fruit stimuli towards themselves.

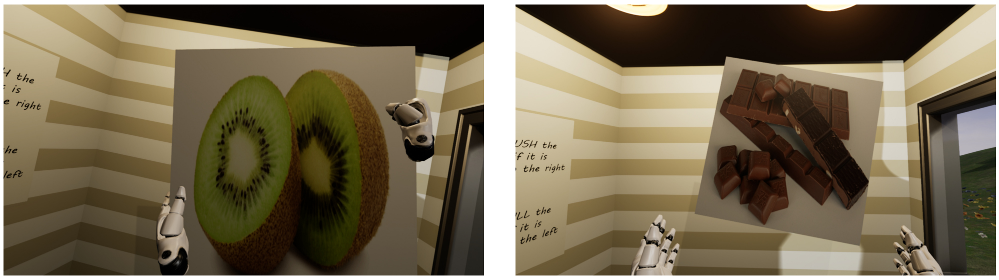
Enjoyment, adapted from Nah et al. (2011), α = .92, (1 = "strongly disagree", 7 = "strongly agree")
I found doing the cognitive task in virtual reality... ...enjoyable. ...interesting. ...boring. ...fun.
Craving Experience Questionnaire - Strength, adapted from May et al. (2014), α = .87 (pre), α = .89 (post), (1 = "not at all", 7 = "extreme")
Right now... ...how much do you want chocolate? ...how much do you need chocolate? ...how strong is the urge to have chocolate? ...how hard are you trying not to think about chocolate? ...how intrusive are your thoughts? ...how hard is/was it to think about anything else? Right now, how vividly do you... ...picture chocolate? ...imagine the taste of chocolate? ...imagine the smell of chocolate? ...imagine what chocolate would feel like in your mouth or throat?
Object Presence Manipulation Check, two items, self-developed, α = .66
How natural was the interaction (gripping, pulling, pushing) with the objects?" (1 = "not natural at all", 7 = "very natural") How strong was your impression that the virtual objects you pushed or pulled were actually in front of you? (1 = "not at all", 7 = "very strong")
Telepresence Manipulation Check, four items adapted from Nah et al. (2011), α = .62, (1 = "strongly disagree" to 7 = "strongly agree")
I could forget the real world around me, as i was moved in the virtual world. When I took of the VR glasses, I felt like I came back to the real world after a journey. During the exercise in the virtual reality, I forgot that I was in the middle of an experiment. The virtual world seemed to me "somewhere I visited" rather than "something I saw."

Table 5.9: Items with Cronbach's α-Values (Paper 2)

High Object Presence (in Approach Condition)

a) Pull Object b) Push Object

Low Object Presence (in Avoid Condition)



c) Pull Image d) Push Image

Figure 5.6: Design Elements for High and Low Object Presence (Paper 2)

5.2.4.4 Procedure

When participants entered the laboratory, they were told that the experiment was about the effects of virtual reality on taste experiences and signed an informed consent form. Afterwards, the experimenter brought them to a desktop computer, where they completed measurements of hunger, noted the time they had last eaten and drunk something other than water, and the CEQ. Next, the experimenter explained how the pre-measure of the AAT on the desktop computer worked and how to use the joystick. After participants had completed the AAT, they were led to another room to conduct the IVR training, which was randomly selected for them.

In the next part of the experiment, the participants completed the presence and enjoyment questionnaires. In the next step, participants completed the chocolate tasting and the post-measure of the AAT at the desktop computer in randomized order. Finally, they completed the post-craving measure and were thanked and debriefed.

5.2.5 Results

We used 2x2 ANOVAs with the training condition and object presence as between-subject factors to check whether the object presence manipulation was successful and to assess the hypotheses related to enjoyment and consumption. Additionally, we used 2x2x2 mixed ANOVAs with the training condition and object presence as between factors and time as a within factor to assess the hypotheses with approach bias as the dependent variable and to gain more insights into the craving effect. The means and standard deviations of the measured constructs are displayed in Table 5.10. The use of 2x2x2 mixed ANOVAs instead of a 2x2 between-subjects ANOVA on the difference scores enables us to assess whether there is a main effect (e.g., did the bias change from pre to post measure?) in addition to the hypothesized interaction effects with time (e.g., did the bias change from pre to post measure differently in the approach and avoid conditions?). Thus, this approach provides more information than an ANOVA on the change scores and facilitates future research on the main effects.

	High		Low	
Construct	Approach	Avoid	Approach	Avoid
1. Enjoyment	5.25 (0.54)	5.08 (0.57)	5.00 (0.35)	4.52 (0.81)
2. CEQ-S (Pre)	3.52 (1.07)	3.41 (1.12)	3.45 (1.14)	3.61 (1.39)
3. CEQ-S (Post)	3.39 (1.38)	3.7 (1.35)	3.04 (1.11)	3.07 (1.24)
4. Chocolate Consumption	15.03 (8.65)	16.01 (12.71)	18.33 (10.69)	20.32 (11.30)
5. Chocolate Bias (Pre)	-46.47 (216.57)	-83.03 (145.58)	-84.8 (114.22)	-63.55 (99.55)
6. Chocolate Bias (Post)	-9.76 (178.17)	-75.95 (104.19)	-67.42 (145.84)	-76.5 (73.63)
7. Fruit Bias (Pre)	-73.03 (143.97)	-76.65 (121.32)	-62.75 (104.57)	-123.73 (179.50)
8. Fruit Bias (Post)	-20.93 (112.69)	-40.42 (129.64)	-81.80 (125.24)	-18.80 (76.73)
9. Object Presence	5.54 (1.05)	5.55 (0.78)	5.02 (1.38)	4.5 (1.78)
10. Telepresence	5.41 (0.81)	4.6 (1.20)	5.12 (0.87)	4.71 (1.21)

Table 5.10: Means and Standard Deviations for Dependent Variables and Manipulation Checks (Paper 2)

Manipulation checks. We used telepresence and object presence as manipulation checks for our conditions. Because both scales refer to a similar construct, we conducted a confirmatory factor analysis with both scales. The results with all items and both scales modelled as separate factors revealed sufficient discriminant and convergent validity according to the Fornell-Larcker criterion. Specifically, the square root of the average variance extracted (AVE) was .60 for telepresence and .79 for object presence, whereas the two constructs correlated with a strength of .59. However, in terms of the factor loadings, item 3 of the telepresence measurement showed only a mediocre loading of

.213, whereas all other items showed values above .50. We retained the item because its content represents an important aspect of telepresence (forgetting that one is part of an experiment), it is part of a scale validated in previous studies, and the significant effect reported below increased in significance after removal of the item. Therefore, removing the item for purely statistical reasons might lead to reduced construct validity and overestimation of significance. Nevertheless, we report the results for both the full and reduced telepresence scales below. First, we assessed whether the object presence manipulation was successful. The ANOVA on perceived object presence showed a significant main effect of object presence ($F(1, 78) = 9.36, p = .014, \eta_G^2 = 0.07$), with higher values for objects than for pictures, indicating that participants using 3D objects perceived higher object presence ($M = 5.50, SD = 0.92$) than participants using 2D pictures ($M = 4.76, SD = 1.60$). Neither the main effect for training condition ($F(1, 78) = 0.52, p = .474, \eta_G^2 = 0.007$) nor the interaction effect for object presence and training condition ($F(2, 78) = 1.14, p = .287, \eta_G^2 = 0.013$) were significant. We further assessed whether the use of 2D pictures or 3D objects changed the perceptions of telepresence. The ANOVA revealed no main effect for the object presence condition (full scale: $F(1, 67) = 0.13, p = .726, \eta_G^2 = 0.002$, reduced scale: $F(1, 67) = 0.16, p = .069, \eta_G^2 = 0.097$). Unexpectedly, a significant main effect for the training condition was found (full scale: $F(1, 67) = 6.19, p = .015, \eta_G^2 = 0.081$, reduced scale: $F(1, 67) = 6.19, p = .009, \eta_G^2 = 0.097$), reflecting that participants in the approach condition (full scale: $M = 5.27, SD = 0.84$, reduced scale: $M = 5.59, SD = 0.92$) experienced a higher level of telepresence than participants in the avoid condition (full scale: $M = 4.66, SD = 1.19$, reduced scale: $M = 4.87, SD = 1.29$). Overall, these results indicated that the object presence manipulation was successful.

Enjoyment. The ANOVA revealed a significant main effect for object presence ($F(1, 79) = 9.36, p = .003, \eta_G^2 = 0.107$), with high object presence leading to higher enjoyment ($M = 5.16, SD = 0.56$) than low object presence ($M = 4.76, SD = 0.67$), supporting Hypothesis 1. Additionally, there was a significant main effect for training condition ($F(1, 79) = 6.19, p = .015, \eta_G^2 = 0.075$), reflecting higher enjoyment for approaching ($M = 5.12, SD = 0.47$) than for avoiding ($M = 4.79, SD = 0.75$) chocolate. The interaction effect was not significant ($F(2, 79) = 1.32, p = .252, \eta_G^2 = 0.017$).

Craving. For craving, an ANCOVA on the post score for the CEQ as a dependent variable and the pre score as a covariate revealed a significant main effect of object presence ($F(1, 79) = 5.80, p = .018, \eta_G^2 = 0.044$) and a significant positive effect of the pre score ($F(1, 79) = 47.44, p < .001, \eta_G^2 = 0.361$). Therefore, post scores in the high object presence condition were higher than those in the low object presence condition, over and above the effect that individuals with higher craving pre scores showed higher post craving scores. Thus, Hypothesis 2 was supported. To further explore whether this difference arose from an increase or decrease in craving over time, we used a mixed 2x2x2 ANOVA. The ANOVA revealed a significant time x object presence interaction effect ($F(1,79) = 5.50, p = .022, \eta^2_G = 0.013$), which is displayed in Figure 5.7. No other main or interaction effects were significant (all ps $> .102$). Separate ANOVAs for high and low object presence revealed that this interaction was due to a main effect of time, reflecting that individuals in the low object presence condition experienced lower craving after the training ($M = 3.06$, $SD = 1.16$) than before the training ($M = 3.53$, 1.26, $F(1,41) = 8.40, p = .006, \eta^2_G = 0.038$). The other main and interaction effects for low and high object presence did not reveal significant differences (all ps $> .225$).

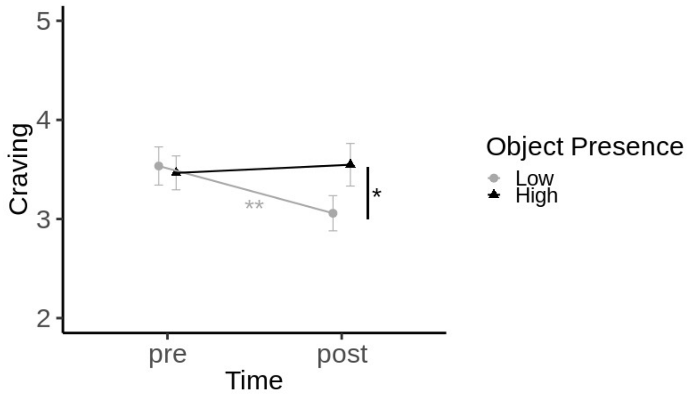


Figure 5.7: Interaction Effect for Craving (Paper 2)

Health-promoting behavior. Because the assumption of normality was violated, we used a logarithmic transformation on the chocolate consumption scores, similar to previous research Schumacher et al., 2016. The ANOVA for chocolate consumption revealed a significant main effect for object presence ($F = 4.00, p = 0.049, eta^2_G = 0.048$), reflecting the result that participants in the low object presence condition ($M = 19.35g, SD = 10.93$) ate more chocolate than those in the high object presence ($M = 15.52g$, $SD = 10.74$) condition. All other main and interaction effects revealed no statistical significance. The results indicate support for Hypothesis 3 but not for Hypothesis 4.

Approach bias. Reaction times (RTs) were recorded from the time the picture appeared until the time the picture disappeared. We excluded error trials (participants pushed or pulled in the wrong direction, 5.7% of all trials) for the approach bias analysis. For the remaining trials, we subtracted the median reaction times using the formula $RT_{Push} - RT_{Pull}$ to calculate approach bias scores (Rinck & Becker, 2007). Thus, positive values for approach bias scores mean that individuals pulled images faster than they pushed them away, whereas negative values indicate faster pushing than pulling.

For chocolate bias, there was neither a main effect nor an interaction (all ps $> .297$). To investigate fruit approach bias, we had to exclude one participant whose reaction time in the post-measure was more than five times above the mean, creating an artificial interaction effect. After this exclusion, only the main effect for time was significant ($F(1, 78) = 5.80, p = .018, \eta_G^2 = 0.027$), indicating that the fruit approach bias scores were higher after training ($M = -39.96, SD = 113.01$) than before training ($M = -85.01, SD = 140.77$), regardless of the condition. Therefore, participants learned to show less avoidance for fruit stimuli, and Hypothesis 5 could not be supported.

5.2.6 Discussion

Regarding enjoyment, our results show that participants enjoyed the AAT version with high object presence more than the AAT version with low object presence, supporting Hypothesis 1. Unexpectedly, individuals also experienced higher enjoyment for the approach than for the avoid training. Supporting Hypothesis 2, the results for craving indicated that whereas chocolate craving scores remained stable for high object presence, participants experienced less craving after using the low object presence AAT version. In line with Hypothesis 3, the high object presence version of the AAT also led to increased health-promoting behavior compared to that of the low object presence version, indicating that training with high object presence is more effective because of the higher similarity (e.g., with regard to craving) to real situations in which the learned behavior has to be recalled. Surprisingly, we found no support that the AAT could reduce chocolate consumption (Hypothesis 4) or bias scores (Hypothesis 5). The theoretical and practical implications of these results are explained below.

5.2.6.1 Implications for Theory

Our study is a first step towards building a theory that explains learning in IVR. Whereas previous studies in psychology have considered IVR as a means to achieve a desirable outcome (Ferrer-Garcia et al., 2015; Ferrer-Garcia et al., 2019), our study highlights the need to investigate specific design elements in IVR to explain why IVR has beneficial learning effects. Furthermore, in contrast to previous IS research that investigated the effects of 3D objects in virtual environments on a desktop computer (Nah et al., 2011), our study shows that the ability of IVR to create possibilities for acting that closely resemble reality can unravel effects in a way that is not possible with traditional technologies. IS research can benefit from further investigating object presence and additional building blocks of IVR to identify specific working mechanisms and boundary conditions in which learning within and outside the health domain occurs in IVR. We discuss specific contributions to the literature on presence, dual process theories, and the identification of design elements in an AAT version for health and learning below.

First, the study contributes to research on presence by identifying the relevance of object presence for health-related experiential and behavioral outcomes. Thus, even though previous research argued that the construct of object presence holds an "apparent lack of relevance to IS research" (Schultze, 2010, p. 437), our results show the opposite. The result that cues with high object presence elicited higher enjoyment (H1) strengthens and refines previous theories in presence research. Whereas telepresence has been shown to be associated with enjoyment in previous research at desktop computers (Nah et al., 2011), our study provides the first evidence that increasing object presence by using 3D objects can be sufficient to increase enjoyment, even when telepresence is held constant. Related to research on motivation for therapy and against the background of technology acceptance models with enjoyment as an antecedent of acceptance (Lowry et al., 2013; van der Heijden, 2004), this finding provides initial evidence that object presence could be a relevant design element to increase motivation and use intentions among patients. In line with this conclusion, the finding that stimuli with high object presence elicited higher chocolate craving after the IVR experience than stimuli with low object presence (H2) further strengthened the proposition that object presence suffices to influence experiential outcomes and shows that this hypothesis even generalizes towards health-related experiential outcomes. This result further clarifies the observed effects of the study of Gorini et al. (2010) by showing that 3D stimuli in IVR do not necessarily need an environment with additional environmental cues to affect health-related experiential outcomes. Furthermore, our study generalizes their results for craving in a nonclinical sample. Finally, the finding that object presence could reduce chocolate consumption (H3) may even extend these findings to actual behavior. However, future research still must investigate whether the effect of object presence for stimuli persists when individuals are present in an environment providing environmental cues (e.g., a restaurant).

Interestingly, in contrast to previous research on the AAT on a desktop computer (Dickson et al., 2016), we did not find an increase in chocolate craving scores from pre- to post-measurement. Whereas the level of craving in the high object presence condition remained stable, craving in the low object presence condition was reduced. This occurred even though both our study and the study of Dickson et al. measured chocolate craving after chocolate tasting. An explanation for this finding could be that the IVR AAT training is associated with a higher level of physical activity than the traditional AAT conducted on a desktop computer. Whereas users of the traditional AAT move a joystick while sitting in front of a computer, participants in our IVR AAT training had to conduct the AAT training while standing and moving their arm to a much larger degree to fulfill the task. Thus, this form of physical exercise could have had diminishing effects on chocolate craving similar to those of other forms of physical exercise, such as a short walk (Ledochowski et al., 2015).

Second, our study contributes to dual process theories by evaluating and comparing the effect of object presence and AAT training on questionnaire and behavioral measurements. The finding that object presence reduced chocolate consumption indicates that the association of the sight of chocolate with eating chocolate was successfully extinguished and implies that new responses became associated with this cue. Importantly, this result that chocolate consumption was reduced while craving remained high a) contributes to the existing literature indicating that chocolate consumption and craving reflect independent processes in the impulsive system (Dickson et al., 2016), b) suggests that learning in a situation with high craving helps to provide a certain immunity to the cue, and c) provides support for the assumption of classical conditioning (van den Akker et al., 2018) that learning in a context with higher similarity to the situation in which the learned content is recalled facilitates learning effects. Furthermore, as we measured craving using a questionnaire, it might be that reflective processes had an additional impact on the craving measure. Thus, whereas individuals might have reflectively come to the conclusion that they want and need chocolate, training the eating-related impulsive system multiple times in the AAT training could have prevented them from enacting behavior on behalf of the reflective system.

With regard to the specific effects of the AAT, the question arises as to why the AAT did not consistently change approach bias (H5) and why we could not find evidence for a reduction in chocolate consumption after the AAT avoidance training (H4). In contrast to our hypothesis, chocolate consumption increased in both avoidance training conditions of high and low object presence, even though this result did not reach statistical significance. Thus, considering that previous studies of the AAT have revealed mixed results with regard to training effectiveness, one explanation for our results could be that the AAT is not suited for changing food consumption. Additionally, we cannot rule out the possibility that training to avoid high-calorie food in the AAT increases chocolate consumption in the form of a behavioral rebound effect, as indicated by Study 3 in Becker et al., 2015. On the other hand, it might be that neither AAT design in IVR was suitable for affecting chocolate consumption with sufficient strength. Whether the problem lies in the specific design of the stimuli in our study or in the exclusion of relevant design elements in IVR is still unclear. Thus, our study highlights the need for future research to investigate different stimuli and additional design elements in IVR, such as the inclusion of environmental cues (e.g., a TV screen with a couch or a laser scan of the room in which food tasting takes place) in the training. Additionally, these findings demonstrate the need for IS research to investigate a) how the different components (e.g., craving, approach bias) of the impulsive system can be measured using different forms of technology such as IVR, b) what differential effects design elements have on the impulsive and reflective systems, and c) how impulsive processes can be measured in real time with NeuroIS methods (vom Brocke et al., 2020) to design adaptive systems that can improve health-promoting behavior.

For the AAT in the eating domain, our study is the first to investigate the design elements of the AAT in IVR. Although object presence could decrease chocolate consumption, AAT avoidance training failed to reveal a significant effect. Considering the specific design variants of the AAT, this could suggest that the specific design for approach and avoidance were still ambiguous for participants. In the traditional AAT version on a desktop computer, visual feedback (i.e., decreasing/increasing picture size when pushing/pulling the picture) is necessary to reduce the ambiguity of the arm movement, and thus the increased complexity of the task in IVR could lead to additional ambiguity. On the one hand, participants must reach towards the object and grasp it before they can push or pull it; therefore, both reactions could be interpreted as an approach. This ambiguity could be reduced by avoiding grabbing when interacting with the stimuli (Schroeder et al., 2016) or by implementing a full-body interaction in which participants step away or towards a stimulus after seeing it (Stins et al., 2011). On the other hand, the repeated interaction with 3D objects (pictures in the case of low object presence and chocolate/fruit objects in the case of high object presence) is closer to reality than the interaction with 2D pictures in the traditional AAT version. Thus, this interaction could be more easily interpreted as avoidance in both the approach and avoid conditions because in both conditions, the stimuli are neither virtually consumed nor permanently stored in direct proximity to the body, even though the IVR environment would allow such a scenario.

Finally, the unexpected result that approach bias scores did not change as hypothesized could be explained by the difference between the training conditions, which took place in IVR, and the bias measurement, which took place on a desktop computer. This implies that even the approach-avoidance-related part of the impulsive system is composed of different, possibly unrelated sub-constructs that need to be measured under conditions with high similarity to the training environment. Whereas previous research has mainly applied AAT training that used the same technology for training and measurement (Dickson et al., 2016; Schumacher et al., 2016), we used IVR for training and a desktop computer version of the AAT for measurement. Thus, regarding implications for dual process theory, it might be the case that the AAT is more sensitive than previously assumed. This would imply that different approach biases are likely to exist that could be measured through different means. This is theoretically plausible given that it is a pattern already shown for different bias measurements, for example, for approach bias in comparison to the implicit association test (Wiers et al., 2016). Specifically, it might make a difference whether approach is measured or trained through different means, for example, a) on a desktop computer, b) in IVR with interaction through controllers, c) in IVR with interaction through virtual hands, d) in IVR with full-body interaction. As a consequence, future research would benefit from comparing different interaction types when measuring and training approach biases. Moreover, this unexpected result represents a basis for IS research, especially NeuroIS research, to further investigate which components the impulsive system consists of, which design elements are necessary to measure them, and which specific design elements are necessary to change them. Overall, our findings inform design research by helping to identify additional ways to increase effectiveness in IVR. Against the background of the mixed results on approach bias scores, a first step for future research would be to test whether high object presence AAT training can change approach bias scores if measurement and training take place with the same IVR AAT version.

5.2.6.2 Implications for Practice

Our study has several implications for practice. First, our study notes that IVR can be successfully used to facilitate health behavior. In this area, the use of 3D stimuli for cue exposure provides an opportunity for training effects that might not be provided by traditional technologies, such as desktop computers, which cannot display 3D objects with sufficient object presence. Therefore, distributing IVR trainings online that aim to improve health-related outcomes may provide an additional positive effect compared to the effect of traditional trainings. In particular, the combination of smartphones with IVR to display 3D objects (e.g., with smartphone headsets) could be an easily accessible technology for most consumers. Moreover, whether displaying 3D objects on a smartphone elicits the same positive effects on health-related behavior is still an open question that could be further investigated. In this regard, the individualization of cues by the user by photogrammetry through a smartphone could be another fruitful road for designers to create cues for which users have a high craving. Second, by showing that the use of 3D objects instead of 2D objects increases training effectiveness, we can conclude that augmented reality and holographic devices can be sufficient in the place of IVR. This can provide opportunities for individuals who cannot use IVR because of motion sickness or because they are too young. Thus, once such devices have become available as consumer technologies, providing online trainings designed for them may be another promising opportunity. Third, our study showed that using 3D objects increased enjoyment of the AAT training over that experienced with 2D images. This indicates that an IVR version of the AAT or cue exposure treatment could facilitate long-term engagement compared to using devices in which it is only possible to display 2D images. This is especially relevant for online trainings, in which motivation can be of greater importance than in offline learning settings. Fourth, the result that whether participants had trained to approach or avoid cues seemed to be irrelevant to health behavior suggests that pushing away objects by touching them is too ambiguous because it might be seen as training initial approach. Therefore, designers should be careful to design avoidance in a way that reduces such ambiguity. This could be done by a full-body approach and avoidance (e.g., stepping forward/backward after a cue appears) or by hand movements that touch the cues for approach but increase physical distance between the cues and the hand for avoidance.

5.2.6.3 Limitations and Outlook

As with all research, our study has some limitations. First, with 83 female participants, our sample size could be higher. It cannot be ruled out that we missed some effects due to low power in relation to H4 and H5, and we could not yet test whether our effects generalize to all genders. Given that we have included effect sizes in all our analyses, future research can use them to conduct studies with higher power. Additionally, it would be interesting to investigate whether gender effects may arise. Apart from possible gender effects in relation to enjoyment, this angle could especially be interesting in relation to the question of whether specific design elements have similar effects on health-related behavior for all genders. Furthermore, it would have been preferable to measure approach bias directly in IVR with the same task as the training, but this was not possible for technical reasons. Third, we assessed chocolate consumption directly after one AAT training session, but the results could have been different with a longitudinal study including more training sessions. Additionally, our conditions differed slightly regarding the virtual room in which the participants were trained (e.g., striped wall in the low object presence condition, presence of a table on which the objects were placed) and the size of the objects. This was done to have an optimal design for both conditions: If pictures would have been presented to be as small as objects, the content of the pictures could easily have been overlooked. Moreover, whereas it was possible to design the low object presence condition similar to AAT versions used on a desktop computer (tilting the pictures to the left or right as an approach/avoidance cue) by using a striped background, this was not possible for high object presence because a rotation of realistically sized 3D food objects can be easily overlooked in an IVR environment. Therefore, a red and blue border was chosen as the cue for approach or avoidance. However, it is unlikely that the differences in the room had any effect because we did not find a significant difference in telepresence measures in our manipulation checks. Finally, we decided to measure craving after the chocolate tasting instead of directly after the IVR experience to avoid contamination of the tasting with the craving measurement. Nevertheless, as our study provides initial insights into AAT design related to object presence in IVR, our results provide a foundation for further development of IVR AAT tasks. Specifically, future research can investigate what role the size of objects in the AAT plays, what other forms of interactions could lead to a higher effectiveness of the AAT, and how different implicit biases can be measured inside and outside of IVR.

5.2.7 Conclusion

The objective of this study was to investigate the effects of the design elements of immersive AAT trainings that can be distributed online on experiential and behavioral health-related outcomes. From the perspective of dual process theories, the literature on classical conditioning, and presence, we investigated object presence as a relevant factor influencing enjoyment, craving, and consumption. The results show that the increased object presence induced by 3D pictures in IVR increases enjoyment and craving, whereas chocolate consumption is reduced. Therefore, the study contributes to current IS research on presence by proposing the development of a learning theory in IVR considering different types of design elements. In line with this, the study shows that accounting for object presence in theoretical explanations of the effectivity of IVR, especially in cases in which traditionally non-immersive tasks are transferred to IVR, helps to transfer learning effects into reality. Furthermore, the study strengthens the incorporation of dual process perspectives into research on online learning systems.

5.3 Paper 3: Gamification Elements in Immersive Virtual Reality

Title	**Gamification Elements in Immersive Virtual Reality. Comparing the Effectiveness of Leaderboards and Co-presence for Motivation**
Author	Katharina Jahn katharina.jahn@uni-siegen.de
Publication Type	Conference Proceedings
Publication Outlet	International Conference on Advanced Collaborative Networks, Systems and Applications
Outlet Information	–
Status	published
Full Citation	Jahn, K. (2019). Gamification elements in immersive virtual reality. Comparing the effectiveness of leaderboards and co-presence for motivation. *Proceedings of the 9th International Conference on Advanced Collaborative Networks, Systems and Applications (COLLA),* 8–11*, Rome, Italy*

Table 5.11: Fact Sheet Paper 3

Gamification Elements in Immersive Virtual Reality. Comparing the Effectiveness of Leaderboards and Copresence for Motivation

Abstract

With the ability to present a completely different environment to users through head-mounted displays, immersive virtual reality (IVR) offers many opportunities to enhance users' motivation and learning. Recent research in the sports context indicates that social facilitation effects occurring with real humans do not necessarily arise when users see a virtual human on a 2D screen. However, whether the increased copresence that immersive virtual reality offers a) can provide increased social facilitation effects compared to 2D screens and b) provides enhanced effectivity compared to traditional gamification elements is still unclear. To investigate this research gap, a 2 (copresence: low vs. high) x 2 (leaderboard: no leaderboard vs. leaderboard) between-subjects laboratory experiment is proposed in this research in progress paper. The expected results can contribute to explain the effects of gamification elements in IVR for intrinsic motivation and performance.

Keywords— virtual reality, copresence, gamification, multi-user, leaderboards.

5.3.1 Introduction

With Immersive Virtual Reality (IVR) technology becoming more and more affordable, new opportunities arise to facilitate learning. IVR has not only the ability to create a high sense of being in a distant environment (telepresence), it can also create a high sense of owning a virtual body (self-presence), being with others (social presence) and being with others in a distant environment (copresence) (Schultze, 2010; Slater & Wilbur, 1997; Witmer & Singer, 1998). The experiences made in IVR can indeed affect cognition and behavior (Falconer et al., 2016; Fox & Bailenson, 2009; Osimo et al., 2015). For example, IVR enables users to see a virtual body visually similar to the self doing sports from both first and third person perspective. When the avatar then gains or loses weight according to activity, long-term activity levels of the user can be facilitated (Fox & Bailenson, 2009). Such designs relying on embodiment of users are not easily possible without IVR.

The characteristics of IVR offer the possibility to design gamification elements used in traditional devices more effectively, especially in relation to learning scenarios with multiple individuals. Gamification describes the use of game elements in non-gaming contexts and requires the use of gamification design elements (Deterding et al., 2011; Liu et al., 2017). Gamification design elements are aimed at motivating or engaging users and are instantiated as objects and mechanics (i.e., interaction rules) (Liu et al., 2017). Related to other virtual individuals, they can consist in the inclusion of leaderboards, e.g., a list of the top ten users or displaying multiple users in the application (Sailer et al., 2017; Schöbel et al., 2017). Whereas the inclusion of gamification design elements, such as leaderboards, satisfies individuals' need to feel competent and might induce increased feelings of autonomy, displaying multiple users can satisfy the need for relatedness and can serve a social facilitation effect (Thiebes et al., 2014). According to research on social facilitation and inhibition effects, being observed by other humans while doing a simple task can create social facilitation, whereas it inhibits task performance for complex tasks (Zajonc, 1965).

For collaborative learning situations, especially the ability of IVR to display quite realistic avatars, which create a high degree of copresence, can create a fundamentally different experience compared to traditional virtual learning environments (e.g., 2D screen at desktop computer). Research on comparing the sense of copresence using a large 2D display or a head-mounted display (HMD) to interact with a single virtual human indicates that individuals can feel the same degree of being colocated in a room with a virtual human in both scenarios (Johnsen & Lok, 2008). However, their perception in which room they were colocated varied, with participants viewing a 2D environment feeling colocated in the actual room, whereas participants with HMD felt colocated in the virtual room. Additionally, it is still unclear how copresence is affected when copresence with multiple individuals should be elicited.

Up to now, whether the higher immersion offered in IVR a) can be used to recreate social facilitation effects present for real humans and b) can compete against traditional gamification elements is still unclear. To address this research gap, this research in progress paper focuses on the area of facilitating engagement in the sports domain in which users located at different places are colocated in a virtual environment and aims at proposing a design methodology to investigate the following research question:

Research Question. Which collaborative gamification design elements lead to increased motivation and performance?

The paper is structured as follows. In Section 5.3.2, the hypotheses are developed on the basis of self-determination theory and literature on gamification. In Section 5.3.3, the

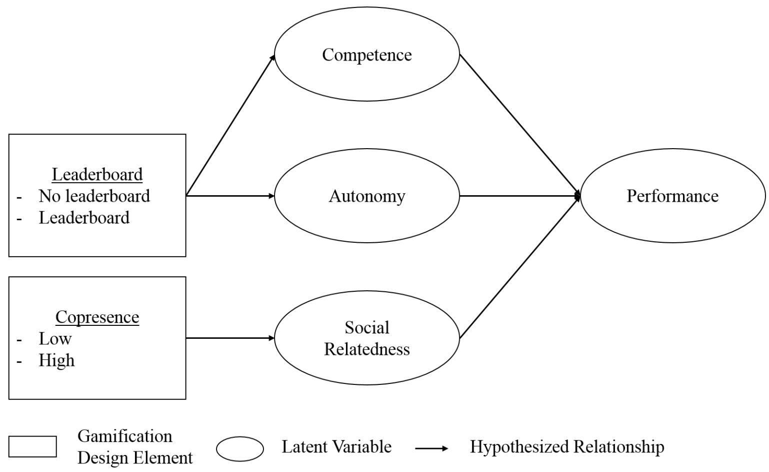


Figure 5.8: Research Model (Paper 3)

methodological approach is described. Finally, Section 5.3.4 concludes with the expected contribution of the proposed experiment and suggestions for future research.

5.3.2 Background and Hypothesis Development

This section describes self-determination theory in relation to gamification to develop hypotheses regarding the effect of copresence and leaderboards on motivation.

5.3.2.1 Self-determination theory and Gamification

Self-determination theory (Deci & Ryan, 1985; Ryan & Deci, 2000) describes how humans develop extrinsic and intrinsic motivation. It proposes that the satisfaction of three psychological needs, competence, autonomy, and relatedness, is relevant for the development of motivation. Need for competence describes that individual strive to experience feelings of achievement during interaction with their environment (Rigby & Ryan, 2011). On the other hand, need for autonomy relates to the experience that actions result from individuals' own volition, whereas need for relatedness describes that individuals strive to belong to other individuals (Deci & Ryan, 1985). The development of the research model (see Figure 1) for this research-in-progress paper on the basis of self-determination theory is described below.

For the area of gamification, self-determination theory can act as a theoretical lens to explain how different gamification elements motivate. Sailer et al. (2017) could show that the inclusion of badges, leaderboards, and a performance graph increased the satisfaction of need for competence and autonomy compared to presenting only points. On the other hand, when users could choose their avatars and are presented with a story, as well as teammates, their need for social relatedness was more satisfied than when they viewed only points. It is therefore hypothesized that the presentation of leaderboards will increase the satisfaction of need for competence and autonomy.

Hypothesis 1. Using leaderboards leads to higher satisfaction of need for competence than using no leaderboards.

Hypothesis 2. Using leaderboards leads to higher satisfaction of need for autonomy than using no leaderboards.

Additionally, increased copresence should lead to higher satisfaction of need of relatedness than low copresence.

Hypothesis 3. High copresence leads to higher satisfaction of need of relatedness than low copresence.

5.3.2.2 Gamification Elements and Performance

Research on the social facilitation effect of virtual humans can be differentiated in whether it has investigated the effects of virtual humans displayed on traditional 2D screens or in IVR with a HMD.

For 2D screens, research has indicated that being with virtual human has similar effects as being with a real human, at least when the task for which performance is measured is a cognitive task. Specifically with regard to inhibition effects, both virtual humans and real humans inhibit performance for female, but not male participants in a pattern recognition and categorization task (Zanbaka et al., 2004). Likewise, with regard to facilitation effects, Liu et al. could show that effects are comparable between virtual humans and real humans, but without detecting gender effects (Liu & Yu, 2018). Additionally, Park et al. could show that social inhibition effects arise for both virtual and real humans in a complex task, whereas for easy tasks, a social facilitation effect could be observed (Park & Catrambone, 2007). However, the social facilitation effect comparing presence versus absence of a virtual human of Park et al. could not be replicated in a recent study (Baldwin et al., 2015). Surprisingly, when the task is not a cognitive task but a sports-related, effects between virtual and real humans become apparent, as shown by a recent study (Snyder et al., 2012). Here, cycling performance could be enhanced when competitive individuals were paired with a real human, but not when they were paired with a virtual human.

In IVR, initial research suggests that social inhibition effects are at a similar level for virtual and real humans, whereas no social facilitation effects could be found for virtual or real humans (Zanbaka et al., 2007). Additionally, computer controlled agents seem to provide less copresence than human-controlled avatars, and here, inhibition effects could only be found for human-controlled avatars (Hoyt et al., 2003). One paper compared the effect of HMD and 2D screens, which indicated that inhibition arises only when using IVR but not when using 2D screens for robotic agents (Emmerich & Masuch, 2016). However, all of these studies were conducted in the domain of cognitive tasks. As the research in progress paper at hand is planned in the context of the sports domain, it can be assumed, in line with research on cycling performance (Snyder et al., 2012), social facilitation effects will arise. However, as Snyder et al. could only find social facilitation effects for individuals paired with a real human, it is hypothesized that the high copresence condition will lead to higher performance than the low copresence condition.

Hypothesis 4. High copresence leads to higher performance than low copresence.

As previous research on gamification elements has shown that leaderboards increase performance (Mekler et al., 2017), the same is assumed for the context of this study.

Hypothesis 5. Using leaderboards leads to higher performance than using no leaderboards.

A meta-analysis in the context of self-determination theory could show that satisfaction of the three psychological needs predicts performance (Cerasoli et al., 2016). We therefore hypothesize:

Hypothesis 6. Performance is positively related to satisfaction of need for competence.

Hypothesis 7. Performance is positively related to satisfaction of need for autonomy.

Hypothesis 8. Performance is positively related to satisfaction of need for relatedness.

5.3.3 Method

In this section, the set-up of the experiment, the gamification design elements, and the planned data analysis is described.

5.3.3.1 Participants and Design

I will use a 2 (copresence: low vs. high) x 2 (leaderboard: no leaderboard vs. leaderboard) between-subjects laboratory experiment with 80 student participants recruited from the local university to test the proposed hypotheses.

5.3.3.2 Materials and Measures

Virtual Reality. Participants will use a virtual environment programmed with Unity 3D displayed with HTC Vive during the experiment. For body tracking, five HTC Vive Trackers (for hip, both feet and both hands) in combination with Hi5 VR Gloves are used. Avatars will be created in Adobe Fuse.

The measurements for the three psychological needs, the manipulation checks, and the indicator for performance are described below. The scales for the three psychological needs and the manipulation checks are measured on a 7-point scale ranging from "strongly disagree" to "strongly agree".

Satisfaction of Need for Competence. The need for competence scale is taken from Sailer et al. (2017) and adapted to the context of the study. The scale consists of four items. One example item is "During the gamified task I had feelings of success".

Satisfaction of Need for Autonomy. The need for relatedness scale is adapted from the autonomy in relation to task meaningfulness scale from Sailer et al. (2017). The scale consists of three items and one example item is "It was worthwile doing the task".

Satisfaction of Need for Relatedness. The need for relatedness scale is adapted from Sailer et al. (2017). The scale consists of three items and one example item is "While doing the task I felt like I was part of a team".

Performance. For learning performance, the times participants have raised their feet in the marching in place task is counted.

Manipulation checks. For copresence, the copresence scale from Poeschl and Doering (2015), as well as the copresence scale from Bailenson et al. (2005) are used as manipulation check, consisting of three items each. An example item is "I was aware that other people were with me in the virtual room." for the Poeschl and Doering scale and "Even when the 'other' was present, I still felt alone in the virtual room" for the Bailenson et al. scale. For Leaderboards, we use the item "'I was informed about how other players performed on the task" as manipulation check.

5.3.3.3 Gamification Design Elements

Leaderboard. Leaderboards will be implemented by displaying the number of repetitions from five other users. In the no leaderboard condition, an empty leaderboard is presented.

Copresence. In the low copresence condition, participants will see four other virtual humans (2 male, 2 female) who will do the task with them on a television screen. On the high copresence condition, the players will be in the same virtual room as the participants.

5.3.3.4 Procedure

One week prior to the first IVR session, we will invite participants to the laboratory to create pictures for the avatars used in the experiment. One week later, when participants enter the laboratory, they will be fitted with the HTC Vive trackers. When they put on the HMD, they will see a room with a large mirror and a large television screen on the wall in front of them. When they look in the mirror, they see the virtual avatar that looks similar to themselves, which they also see from first person perspective. Participants will see their own points above the mirror and television screen. Additionally, participants in the leaderboard condition see a leaderboard displayed above their points, which they will be made aware of by the experimenter.

For participants in the low copresence condition, the television screen will show four participants who enter the room on the screen and train with them. On the other hand, participants in the high co- presence condition, the four trainees enter the same virtual room as the participant. Then, they will be instructed on how to do the marching task. In this task, they have to alternately lift their feet to a specific height displayed in IVR for ten training trials in which the experimenter validates that the participants perform the action correctly. Then, they are told that they can do as many repetitions as they want. After they have finished, participants finish the motivation and presence questionnaire in IVR. Afterwards, they are thanked and debriefed.

Data Analysis: The data will be analyzed using four 2x2 ANOVAs for the three psychological needs competence, autonomy, and social relatedness, as well as performance. Additionally, the complete model will be tested using covariance-based structural equation modeling.

5.3.4 Conclusion and Future Research

The proposed experiment can contribute to literature on gamification and IVR and answering the research question in several ways. First, the study helps to gain insight into which gamification elements are most effective in IVR to increase motivation and performance. Additionally, the experiment contributes to explain motivational working mechanisms of gamification elements against the background of self determination theory. Finally, the experiment can contribute to explain conditions under which social facilitation effects arise. On this basis, future research can investigate whether the proposed working mechanisms of this model generalize to other areas in the sport domain, as well as sport-unrelated domains, such as knowledge work, and application areas outside of IVR. From a practice perspective, collaborative gamification elements can then be used to enhance motivation in multi-user scenarios (e.g., applications supporting health behavior). Furthermore, future research can develop algorithms that implement these collaborative gamification elements efficiently.

5.4 Paper 4: Gamified Helping

Title	**Gamified Helping? The Impact of Individualized and Group-Level Cooperative Evaluation on Knowledge Sharing**
Authors	Katharina Jahn katharina.jahn@uni-siegen.de Bastian Kordyaka bastian.kordyaka@uni-siegen.de Tobias Scholz tobias.scholz@uni-siegen.de Bjoern Niehaves niehaves.bjoern@uni-siegen.de
Publication Type	Conference Proceedings
Publication Outlet	Internationale Tagung Wirtschaftsinformatik
Outlet Information	JOURQUAL3: C
Status	published
Full Citation	Jahn, K., Kordyaka, B., Scholz, T., & Niehaves, B. (2020). Gamified helping? The impact of individualized and group-level cooperative evaluation on knowledge sharing. *Proceedings of the 15th Wirtschaftsinformatik 2020 (WI2020), Potsdam, Germany*

Table 5.12: Fact Sheet Paper 4

Gamified Helping? The Impact of Individualized and Group-Level Cooperative Evaluation on Knowledge Sharing

Abstract

Even though knowledge sharing is an important part of cooperation in organizations to ensure productivity and economic functioning, employees sometimes refrain from sharing knowledge with colleagues. Accompanied by the digital disruption of work, one fruitful way to improve organizational knowledge sharing between employees is the design of technology. However, which technological design elements improve knowledge sharing is still not well understood. This research-in-progress paper combines research on knowledge sharing with the common ingroup identity model to investigate the effect of individualized and group-level gamification design of cooperative evaluations on knowledge sharing behavior. We hypothesize that using group-level design facilitates intrinsic motivation for knowledge sharing by highlighting a superordinate identity. On the other hand, individualized design should enhance knowledge sharing by increasing extrinsic motivation. We plan to test these hypotheses using a 2 (individualized evaluation: badges vs. none) x 2 (group-level evaluation: badges vs. none) between-subjects experiment.

Keywords— gamification, prosocial behavior, cooperation, experiment

5.4.1 Introduction

Knowledge sharing (KS) is an essential part of organizational work that can increase employees' organizational performance and productivity over time (Hooff & Ridder, 2004; van den Hooff & De Leeuw van Weenen, 2004). Unfortunately, employees differ in their motivation to share knowledge according to contextual and individual differences (Constant et al., 1994; Fehrenbacher, 2017; Hsu et al., 2007). As a consequence, Information Systems research can provide pursuing insights into how technology can be designed to increase individual motivation and subsequently KS behavior (Alavi & Leidner, 2001). Gamification, as the use of game elements in non-gaming contexts (Deterding et al., 2011), has shown to affect motivation. A range of different gamification elements exists that are proposed to address motivation to different degrees (Schöbel et al., 2017). Additionally, research has shown that gamification design aimed at cooperation can enhance KS intention and behavior (Morschheuser et al., 2019; Morschheuser et al., 2017). In this research in progress paper, we focus on the design of individualized and group-based evaluative aspects that are connected to displaying badges as rewards for enacting KS behavior. Therefore, we have the goal to foster KS by designing software for mediated communication and aim to investigate the following research question: *How can the individualized and group-level design of gamification elements facilitate knowledge sharing?* To investigate this research question, we plan to conduct a 2 (individualized evaluation: badges vs. none) x 2 (group-level evaluation: badges vs. none) between-subjects experiment aimed at getting more insights into the initial stage of introducing gamification design elements in an organization.

5.4.2 Theoretical Background and Hypothesis Development

The research model is displayed in Figure 1 and the hypotheses are developed in relation to theory below.

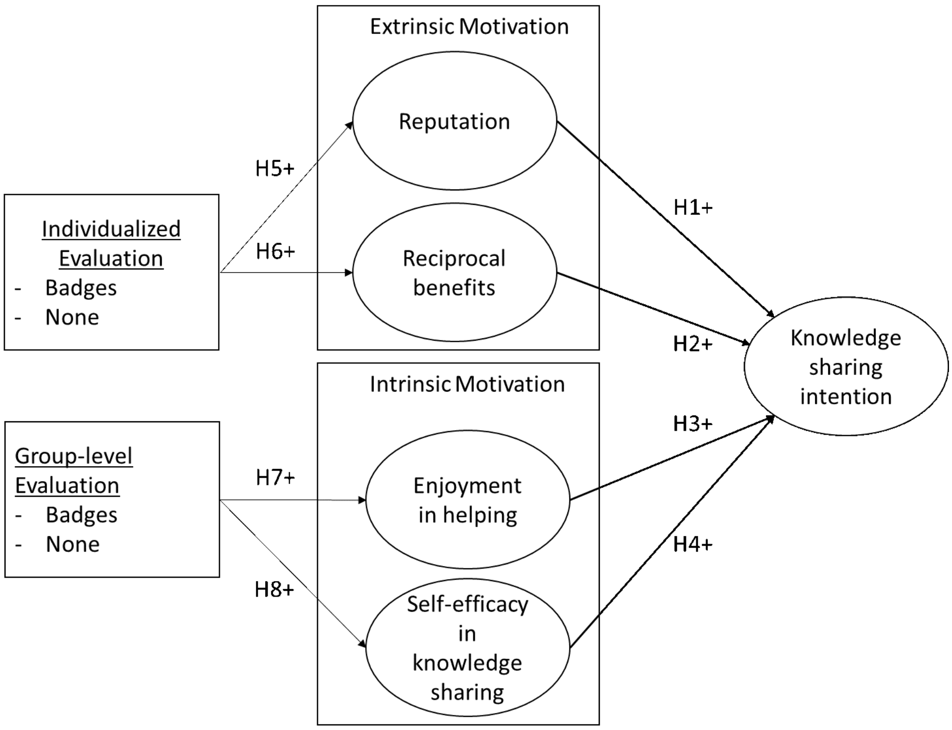


Figure 5.9: Research Model (Paper 4)

5.4.2.1 Knowledge Sharing

Knowledge can be defined as consisting of "information, ideas, and expertise relevant for tasks performed by individuals, teams, work units, and the organization as a whole" (Bartol & Srivastava, 2002, p. 65). As a consequence, KS describes the process of contributing knowledge to other individuals or groups (Bock et al., 2005; Hsu et al., 2007). KS has been associated with a range of predictors, including environmental factors, individual characteristics, and motivational factors (Wang & Noe, 2010). From a motivational perspective, KS can occur out of extrinsic and/or intrinsic motivation (Rode, 2016), which has already been investigated in the context of different KS applications (Kankanhalli et al., 2005; Lin, 2007; Rode, 2016). This research showed that extrinsic predictors of KS consist of reputation and reciprocal benefits, whereas intrinsic predictors are self-efficacy and enjoyment (Kankanhalli et al., 2005; Lin, 2007; Rode, 2016). However, evidence for enjoyment as predictor is still mixed (Rode, 2016). On the basis of this literature, we hypothesize that perceived reputation, reciprocal benefits, enjoyment in helping others, and self-efficacy in KS leads to increased KS intention (H1-4) to gain more insights into motivationally relevant predictors for KS.

5.4.2.2 Gamified Design

Extrinsic motivation: Reputation and reciprocal benefits constitute indicators for extrinsic motivation. Reputation describes the evaluation of a person based on their past actions, whereas reciprocal benefits refer to the idea that someone else or the helper benefits after someone has received help (Baker & Bulkley, 2014). As a consequence, reputation is increased when KS activities are visible to other individuals in the organization (Rode, 2016). In line with this, expectations for reciprocal benefits should also be increased if evaluations of past cooperative behavior is disclosed. We therefore hypothesize that providing individualized evaluation on KS activities increases expected reputational benefits (H5) and reciprocal benefits (H6).

Intrinsic motivation: Enjoyment in helping others and self-efficacy in KS are indicators of intrinsic motivation. Enjoyment in helping others describes a positive feeling that is related to enacting in altruistic behavior (Lin, 2007). Based on the common ingroup identity model (Gaertner et al., 1993; Gaertner et al., 1989) individuals should engage in more altruistic behavior if a superordinate group category is made salient. In the case of KS, group-level evaluation consists of such a superordinate group category. Accordingly, when group-level evaluation is given, individuals should perceive higher enjoyment than when no group-level evaluation is given because they consider themselves part of the group that receives help (H7). On the other hand, self-efficacy describes the perception that one's KS contribution is successful (Hsu et al., 2007). Therefore, we expect that providing group-level evaluation will increase the salience of group-level benefits through KS and subsequently increases self-efficacy in KS (H8) by making possible benefits more transparent.

5.4.3 Method

Participants & Design: We plan to recruit 128 participants through the crowdsourcing platform clickworker.com. Participants will be told that they will participate in a survey investigating the effect of e-mails on motivation. To test the hypotheses, we will use a 2 (individualized evaluation: badges vs. none) x 2 (group-level evaluation: badges vs. none) between subjects design.

Procedure: The procedure related to the non-gamified aspects of the study is adapted from Constant et al. (1994) and Fehrenbacher (2017) and consists of four steps. First, participants are introduced to the gamification system in their company. For this, participants are told that they are junior-level programmers. Additionally, their company has just implemented a system that evaluates employees on the basis of their intraorganizational cooperation. The system consists of five stars and with each cooperative task the amount of yellow in the stars increases. According to evaluation condition, participants are told that individualized and group-level evaluation are [not] used in their working group. Second, participants are introduced to the KS context. They are told that they work with a colleague named Alex in the same department and it is revealed that Alex refused to fix a program bug about a month ago. Third, they see a screenshot of a message in which Alex asks them for help (see Figure 2). The screenshot image adopted from the Radboud Face Inventory (Langner et al., 2010). Fourth, they will answer questionnaires for the dependent variables.

Measures: KS intention will be measured with a question adapted from Fehrenbacher: "What is the likelihood you would give a copy of the program to Alex?" (1 = not at all likely to 7 = very likely). The predictors reputation, enjoyment and self-efficacy are adapted from Kankanhalli et al. (2005) and the reciprocal benefits scale is adapted from Lin (2007). The scales are measured with four items each and were rephrased to the situational context of this study. For all of the following items, participants answered on a 7-point likert scale (1 = strongly disagree, 7 = strongly agree). An example item for reputation is "Sharing my program with Alex improves my image within the organization". An example item for reciprocal benefits is "When I share the program with Alex, I expect to receive knowledge in return when necessary". An example item for enjoyment in helping others is "I enjoy sharing the program with Alex". Finally, an example item for self-efficacy in KS is "I have the expertise needed to provide valuable knowledge for my working group".

Gamification Design: The different gamification designs are displayed in Figure 2. Participants are told that the evaluation was introduced today and therefore everyone has

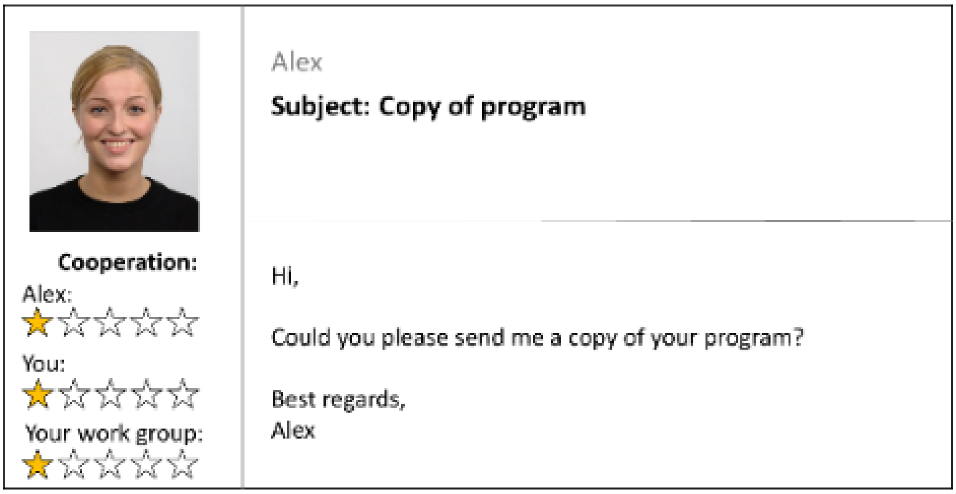


Figure 5.10: Example of Gamification Design (individualized and group-level evaluation (Paper 4)

only the lowest level (one star). By engaging in KS behavior, the level can be increased. For the individualized evaluation design, stars are displayed as badges for the participants (You) and Alex in the badges condition, which indicates their degree of cooperation. In the condition without badges, the area is blank. For the group-level evaluation design, the stars are the same as in the individualized evaluation design, the only difference being that the stars represent the overall degree of cooperation inside the working group as whole.

5.4.4 Outlook

Our next steps consist of programming the questionnaire and pre-registering the study with a detailed data analysis plan and all relevant methodological details (e.g. on https://osf.io). Next, we will start data acquisition and afterwards analyze the data. We expect that the results of our study will contribute to research on KS by identifying the working mechanisms of individualized and group-level gamification designs. Additionally, the results of the study can be used to increase the understanding of KS in relation to gamification elements in organizations. On the basis of our findings, future research can investigate additional gamification designs that facilitate KS through creating a common ingroup. Additionally, the effects of the gamification designs for group coherence and employee conflict can be explored.

6 Track 2: Virtual Bodies

6.1 Paper 5: Between Anthropomorphism, Trust, and the Uncanny Valley

Title	**Between Anthropomorphism, Trust, and the Uncanny Valley: A Dual-Processing Perspective on Perceived Trustworthiness and Its Mediating Effects on Use Intentions of Social Robots**
Authors	Anika Nissen anika.nissen@uni-due.de Katharina Jahn katharina.jahn@uni-siegen.de
Publication Type	Conference Proceedings
Publication Outlet	Hawaii International Conference on System Sciences
Outlet Information	JOURQUAL3: C
Status	published
Full Citation	Nissen, A., & Jahn, K. (2021). Between anthropomorphism, trust, and the uncanny valley: A dual-processing perspective on perceived trustworthiness and its mediating effects on use intentions of social robots. *Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS-54), Maui, Hawaii, USA*, 360–369

Table 6.1: Fact Sheet Paper 5

Between Anthropomorphism, Trust, and the Uncanny Valley: A Dual-Processing Perspective on Perceived Trustworthiness and Its Mediating Effects on Use Intentions of Social Robots

Abstract

Designing social robots with the aim to increase their acceptance is crucial for the success of their implementation. However, even though increasing anthropomorphism is often seen as a promising way to achieve this goal, the uncanny valley effect proposes that anthropomorphism can be detrimental to acceptance unless robots are almost indistinguishable from humans. Against this background, we use a dual processing theory approach to investigate whether an uncanny valley of perceived trustworthiness (PT) can be observed for social robots and how this effect differs between the intuitive and deliberate reasoning system. The results of an experiment with four conditions and 227 participants provide support for the uncanny valley effect. Furthermore, mediation analyses suggested that use intention decreases through both reduced intuitive and deliberate PT for medium levels of anthropomorphism. However, for high levels of anthropomorphism (indistinguishable from real human), only intuitive PT determined use intention. Consequently, our results indicate both advantages and pitfalls of anthropomorphic design.

6.1.1 Introduction

In almost all situations of our lives, first impressions are made in the blink of an eye (Bergmann et al., 2012; Paetzel et al., 2020) and often already predict our further attitude and behavior. The reason for this can be found in first impressions, especially of visual beauty, leading to a halo effect due to which further assumptions about the trustworthiness, warmth, and competence of a robot are made (Bergmann et al., 2012; Paetzel et al., 2020). Further, as social robots are designed increasingly similar to actual humans, anthropomorphism has shown to significantly correlate with evaluations of perceived trustworthiness (PT) (Sanders et al., 2011; Zlotowski et al., 2016), a crucial predictor for use intentions (Gefen et al., 2003; Mathur & Reichling, 2009). Moreover, positive affect towards robots, such as warmth and PT, are pivotal for humans to accept and adopt social robots in their life (Hancock et al., 2011), which is a necessary step to enable comfortable, social human-robot interactions. Especially for the interaction with robots, adding emotional and social interactions tends to reduce the perceived stress and thus, increase PT in the robot with which the interaction took place (Lohani et al., 2016). However, while these aspects are also factors for increased anthropomorphism of robots, the correlation between anthropomorphism and positive affect towards robots does not follow a linear line but enters at a specific level an uncanny valley (Mori, 1970). In the uncanny valley, human actors have increased negative attitudes towards robots, which become positive again when anthropomorphism is almost indistinguishable from a real human (Mori, 1970). Even though the first introduction of the uncanny valley effect happened a century ago and the levels of anthropomorphism in robots have significantly increased since then (see i.e. the robot Erica), the uncanny valley effect seems to still hold true for higher anthropomorphic robots (Beiboer & Sandoval, 2019; Strait et al., 2019).

While increased anthropomorphism has several positive effects, it might also facilitate humans to apply social reasoning towards robots (such as theory of mind). As a consequence, human users may cease to distinguish between humans and robots even though it would be necessary (Culley & Madhavan, 2013). In this respect, the differentiation between two systems of social processing is crucial: (1) an intuitive, affective system, and (2) a cognitive, reflective system (Lobato et al., 2013; Strack & Deutsch, 2004). Given the increasing levels of anthropomorphism in robots, the intuitive system might not be able anymore to make a distinction between human and robot, while i.e. the cognitive system might then detect the processing error (Saygin, A.P. et al., 2010; Wang & Quadflieg, 2014). Given that PT is a complex construct which also consists of emotional and cognitive reasoning (Gefen et al., 2003; Komiak & Benbasat, 2006), the dual-processing theory might therefore be transferable when considering PT as pivotal impact factor for social human-robot interactions (Stoltz & Lizardo, 2018). However, to the best of our knowledge, this perspective has not yet been taken to investigate the mediating effects of anthropomorphism level, intuitive and deliberate PT on further use intentions.

Consequently, this paper aims to investigate, a) *if and how the uncanny valley of PT for human-like robots differs between intuitive and deliberate PT* and b) *how these PT types influence the use intentions of the robot.* To address this research goal, we first review literature regarding PT as a construct and how it can be understood in correspondence to a dual-processing theory. After that, we focus on PT specifically in human-robot interactions by taking robot anthropomorphism as one major influence factor on pre-interaction PT evaluations. Next, we design a study focusing on intuitive and deliberate pre-interaction PT evaluations of social robots on four different levels of anthropomorphism and investigate their mediating effects on use in form of interaction intentions. From the results of this study, we are able to sketch uncanny valleys for the intuitive and deliberate evaluations each, and identify the impact of both reasoning systems.

6.1.2 Related Literature

The construction and introduction of social robots receives more and more attention in various application fields such as education, support for decision makers, healthcare, therapy, at the workplace, or at home (Gockley et al., 2005; Kim et al., 2013; Leite et al., 2013; Lohse et al., 2008). In all of these application fields, the cooperation and collaboration with social robots is crucial for their successful implementation. Therefore, when designing robots for social interactions, there are several (unwritten) rules and norms which robots need to adhere to when they should be accepted in everyday life (Gockley et al., 2005; Nakauchi & Simmons, 2002). For instance, robots require human spatial skills for moving naturally between humans (Kirby et al., 2010). While the degree of fulfilling social requirements is also severely influenced by how the robot behaves, its visual appearance already gives first cues which lead to expectations about its behavior (Leite et al., 2013). That is, solely depending on the first impression of a robot's visual appearance, assumptions about its capabilities and roles in social contexts are automatically made (DiSalvo et al., 2002; Hayashi et al., 2010; Lohse et al., 2008). However, designing for high visual anthropomorphism might not always have the desired effect, since it might lead to expectations about the robot which cannot be met during interaction (Duffy, 2003). Therefore, a focused investigation of the especially relevant construct of PT for high anthropomorphic robots seems reasonable.

6.1.2.1 Dual-Processing Theory and Perceived Trustworthiness

In social interactions, trust and PT are crucial and complex constructs which can be subdivided into different types. One approach from Information Systems research is the distinction between (1) trust beliefs, (2) trusting intention, and (3) disposition to trust (Gefen et al., 2003; Komiak & Benbasat, 2006). In this paper, we further focus on (1) trusting beliefs which are elicited by a robot's visual design features (in means of anthropomorphism) and which are in this paper referred to as PT. PT can be further subdivided i.e. into emotional vs. cognitive PT (Stoltz & Lizardo, 2018), which might also be integrated with one another (Komiak & Benbasat, 2003).

Since PT is a social construct, this integration might be similar to the dual-processing theory of social reasoning, in which system 1 is characterized as affective, automatic, and intuitive processing, while system 2 is thought to be cognitive, rational, and deliberate processing (Lobato et al., 2013). Consequently, the often conceptualized emotional and cognitive component of PT might also be seen as system 1 and system 2 reasoning (Stoltz & Lizardo, 2018). System 1, or in this paper further referred to as "intuitive PT" evaluates stimuli fast and is more prone to erroneous decisions than system 2

(Lobato et al., 2013); however, it can not completely inhibit system 1 (Strack & Deutsch, 2004).

By taking this dual-processing perspective, we make the implicit assumption that human-robot interactions require social reasoning and cognition. Apparently, while robots become more human-like, there is the belief that the social processing of robots will also become closer to that of a human (Lobato et al., 2013). Research stating that humans do not differentiate much between robots and humans when both are perceived as trustworthy (Jessup et al., 2020) further supports this assumption. However, when it comes to how PT is evaluated, there are general differences between how we evaluate the PT of a human and how that of a technology (Lankton et al., 2015). This further implies the question, with which criteria the PT of a robot might be evaluated. Do users assess criteria closer to human-related factors, such as competence or benevolence, or are they more concerned about technological factors, such as reliability and helpfulness? Especially in relation to different levels of anthropomorphism and a possible uncanny valley, these questions might further help to gain deeper insights into how humans perceive and evaluate social robots. Thus, we take these factors into account for measuring deliberate PT, but will not discuss them further.

6.1.2.2 Anthropomorphism and the Uncanny Valley

Studies investigating the antecedents of human PT in robots have identified that both human characteristics and ability, as well as robotic attributes and performance impact how trustworthy a robot is perceived (Hancock et al., 2011; Sanders et al., 2011). For instance, the personality traits of humans can significantly impact their attitudes towards the robot (Elson et al., 2020). Among robot characteristics or attributes, anthropomorphism has also shown to significantly correlate with PT ratings (Sanders et al., 2011) and generally can be represented through visual cues, auditory, or behavioral characteristics of robots (Pfeuffer et al., 2019). In the frame of this work, however, we focus on the visual cues only, as this is about the first cue we perceive and evaluate from a robot, which may lead to starting a conversation and subsequently evaluating its speech and behavior.

When talking about anthropomorphism of robots, however, the uncanny valley effect also needs to be addressed (Mori, 1970). This effect is one of the most relevant approaches to explain how individuals differentiate between humans and robots, or more precisely, between different levels of anthropomorphism (Mathur & Reichling, 2009; Mori, 1970). This effect proposes that if anthropomorphism reaches a certain level between high and low human-likeness, the *uncanny valley* is entered (Mathur et al., 2020), which results in more negative reactions against this entity. Only when anthropomorphism is high and becomes almost indistinguishable from a human, the impression becomes more positive again. While changes in the evaluation of robots between first impression and first interaction are likely (Seymour et al., 2017), we will further focus on first impression and its uncanny valley, only. Further, since Mathur and Reichling (2016) have already shown for a variety of robot faces that the uncanny valley effect generally holds true when affect and PT are evaluated in first impressions, we further explicitly consider PT as dual-process and investigate the differences between intuitive and deliberate PT evaluations not only in regard to an uncanny valley, but also regarding their mediating effects on use/interaction intentions.

6.1.3 Method

6.1.3.1 Sample

We used the online platform clickworker to recruit participants for the survey. 253 participants completed the questionnaire. After excluding participants who were faster than 90% of the sample, to remove participants who merely clicked through the questionnaire without answering the questions seriously, 227 participants remained. 60.79% of participants were male and 38.76% were female and one participant was non-binary/diverse (0.44%). Additionally, participants were between 18 and 69 years old (M = 37.83, SD = 11.74). The majority of participants indicated to be working in the services sector (17.62%), followed by the IT sector (13.22%) and business administration sector (9.69%) The remaining participants (59.47%) worked in a diverse set of other sectors or were currently unemployed.

6.1.3.2 Stimuli

In order to investigate possible uncanny valleys existing for intuitive and deliberate PT, we included robots at four different anthropomorphism levels in our study. Since women tend to be perceived as being more trustworthy (Riedl et al., 2010), we decided to include mainly female robots. More precisely, we included Nao as representative control group for low anthropomorphism which is at the same time the only gender-less robot condition. For higher degrees of anthropomorphism, we used Sophia for medium-low anthropomorphism, Mark1 for medium-high anthropomorphism, and finally, a human for the high anthropomorphism (Figure 6.1). At this point, it needs to be noted that participants were told that all photographs illustrated robots, including the human.

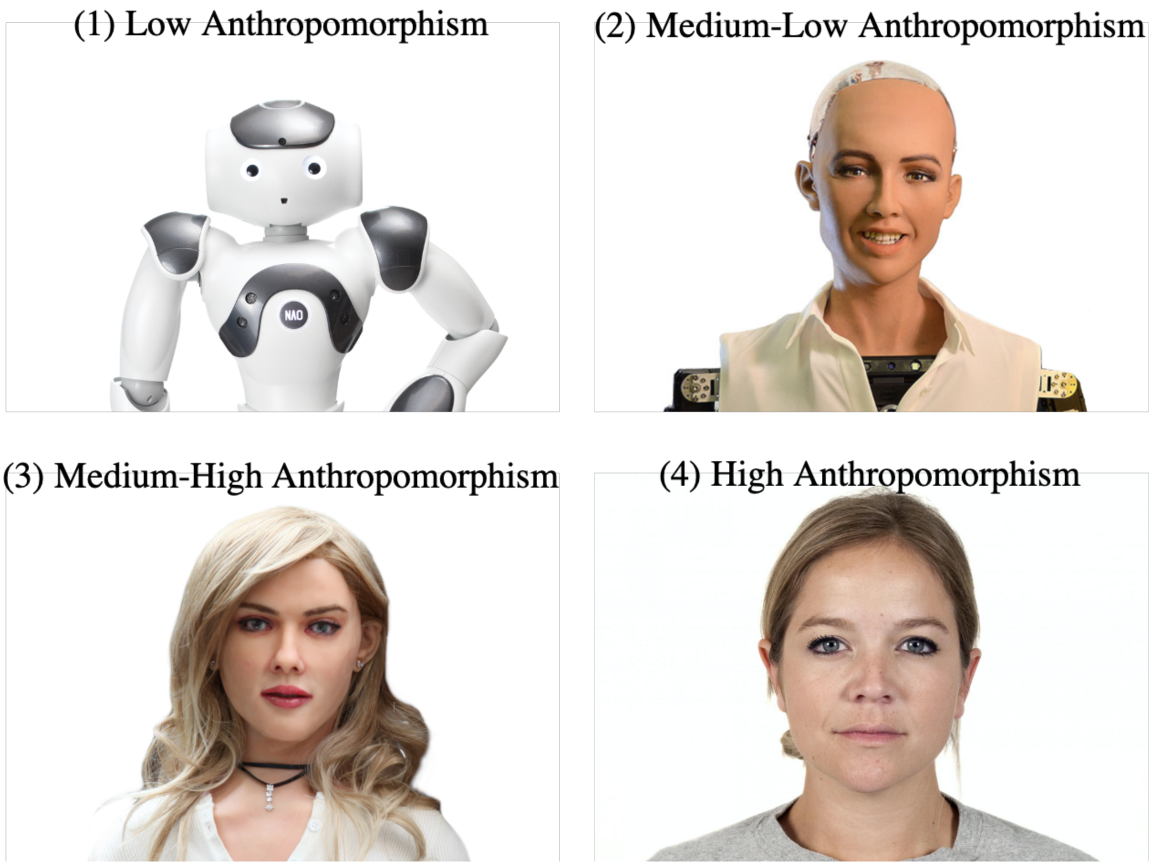


Figure 6.1: Stimuli Used in the Survey (Paper 5)

6.1.3.3 Study Design

To investigate how intuitive and deliberate PT are evaluated based on a robot's visual appearance, we created an online survey designed as follows. First, participants had to rate their general attitude towards technology and their PT perceptions regarding technology, innovations, and humans. After that, pictures of the robots were shown and intuitive PT was measured with time restrictions for the decision to make. This was followed by demographic questions, in which a control question was included to filter out inattentive participants. In case the control question was answered wrong, the questionnaire was closed and participants could not continue. If the control question was answered correctly, participants proceeded to rate the robot conditions for deliberate PT for which the pictorial stimuli were shown and had to imagine they had the opportunity to use the robot in a shopping situation before the scales appeared. Finally, participants answered the humanness manipulation check, after which they were thanked, debriefed and received their clickworker code. This study design is justified by the visual robot appearance having a significant impact on trust evaluations, which is already formed at the first impression of the robot and significantly impacts further evaluations of the robot (Bergmann et al., 2012; Paetzel et al., 2020). Therefore, we use images of robots as stimuli which provide us a first indication of how PT of the robots is evaluated.

6.1.3.4 Measurements

We included a single item scale for intuitive PT that had to be rated on a 5-point Likert scale ("To which degree would you trust this robot?", from 5=completely to 1=not at all). To avoid biases due to the single item, we repeated this question three times for each included robot (i.e., participants completed 12 trials in total) in randomized order and averaged the answers across the three trials for each robot for further analyses. Deliberate PT was measured primarily with the three items for trusting intentions scale from McKnight et al. (2002a) used by Lankton et al. (2015) (e.g., "I can always rely on this robot for buying new products"). Since this construct might not fully capture deliberate PT, we additionally included the PT scales (Lankton et al., 2015) of human-PT being integrity, competence, and benevolence (McKnight et al., 2002a) and of technology-PT being functionality, helpfulness, and reliability (McKnight et al., 2011). All of these scales, however, were only shown for one randomly selected robot (leading to 53 data points for low, 56 for medium-low, and 59 each for medium-high and high condition), while the perceived trusting intentions scale was shown for every robot so that the questionnaire did not get too tiring. Finally, we included a use intention scale for every robot, which was adapted from Davis et al. (1989) (e.g., "I would use this robot to assist me in my buying decision") to further interpret how intuitive and deliberate PT influence use intentions and thus, further HRI. Cronbach's Alpha indicated sufficient reliability for all three scales (intuitive PT: .87, between .94 and .97 for individual robots; deliberate PT: .94, between .91 to .93 for individual robots; use intention: .95, between .93 and .94 for individual robots) (Blanz, 2015). The manipulation check for humanness ("Please indicate to which degree the robots pictured below look like a machine or a person to you") consisted of rating the humanness of each stimuli on a scale from 0% (= machine-like) to 100% (= person-like).

To ensure that intuitive and deliberate processing were primarily measured, we used two means: (1) a time pressure/time delay component and (2) one-item question for intuitive trustworthiness, and a multi-item questionnaire for the deliberate trustworthiness. Using a time component in a questionnaire to distinguish between intuitive and reflective system has shown to be a common method in other studies (Betsch & Kunz, 2008; Glöckner & Witteman, 2009; Sarmany-Schuller, 2010). Therefore, image and scale for the intuitive PT scale was shown for merely 4 seconds in which participants had to make a decision. In case they did not make an input within the 4 seconds, the system remained at the question and robot. For the deliberate PT, first, the image of the robot was shown alone and after 5 seconds the scales appeared in addition to the image and were then clickable. Through this, we tried to ensure that participants looked at the stimuli for a certain time before evaluating the deliberate PT scales taken from McKnight et al. (2011), McKnight et al. (2002a).

	Trust Intention	Integrity	Benevolence	Competence	Reliability	Helpfulness
Integrity	0.79					
Benevolence	0.80	0.83				
Competence	0.76	0.79	0.79			
Reliability	0.78	0.82	0.75	0.80		
Helpfulness	0.80	0.83	0.84	0.90	0.84	
Functionality	0.80	0.82	0.82	0.88	0.82	0.90

Table 6.2: Correlations between PT Scales (Paper 5)

6.1.4 Results

Means and standard deviations are given in Table 6.3. Because the assumption of sphericity was violated for intuitive and deliberate PT, we analyzed the data using oneway repeated measures ANOVAs with Greenhouse-Geisser correction. Additionally, we used Tukey-corrected post-hoc tests for follow-up analyses.

Anthropomorphism	Low M (SD)	Medium-Low M (SD)	Medium-High M (SD)	High/Human M (SD)
Intuitive PT	2.94 (1.08)	2.27 (0.91)	2.53 (1.03)	3.70 (1.03)
Deliberate PT	3.86 (1.51)	3.13 (1.47)	3.34 (1.51)	4.30 (1.53)
Use Intention	4.02 (1.60)	3.25 (1.59)	3.51 (1.54)	4.48 (1.56)

Table 6.3: Means and Standard Deviations for Intuitive and Deliberate PT (Paper 5)

Because we measured only PT intention in a within design, we chose to use it as our main outcome for overall PT. To validate that it covers different facets of PT, we looked at bivariate correlations. All PT scales were significantly correlated, with a minimum value of .75 and the minimum value for PT intention was at .76. Therefore, the measures were highly correlated which lets us assume that the trust intention scale sufficiently represents the different PT facets as a measure.

6.1.4.1 Manipulation Checks

The ANOVA for the manipulation check of anthropomorphism was significant ($F(2.52, 559.74) = 1106.98, p < .001, \eta_G = .735$). Post-hoc tests showed that the high condition was seen as most human-like (94.97%), followed by the medium-high condition (54.15%), the medium-low condition (30.43%), and then the low condition (7.62%) (all with $p < .001$, on a scale from 0% being machine-like and 100% being person-like). Therefore, the anthropomorphism manipulation was successful.

6.1.4.2 Experimental Results

Intuitive PT: The repeated measures ANOVA revealed a significant main effect ($F(2.64, 597.27) = 123.56$, $p < .001$) pointing to an uncanny valley effect. That is, post-hoc tests showed that this was due to higher PT ratings for the high condition than for any other condition, followed by the low anthropomorphism condition, the medium-high anthropomorphism condition, and, finally, the medium-low condition. All *p*s were < .001 except for the difference between the medium-low and medium-high conditions, which was at .007. Consequently, these results support the uncanny valley effect for the intuitive PT for which current high anthropomorphism levels still do not seem to be high enough.

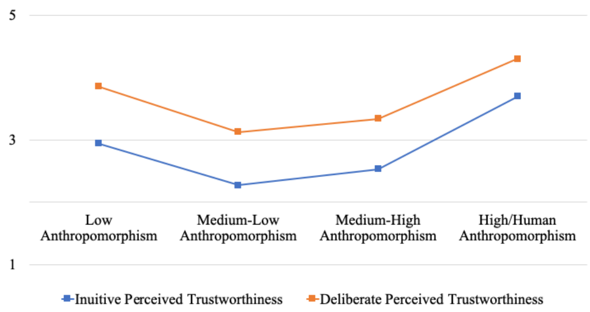


Figure 6.2: Uncanny Valleys of Intuitive and Deliberate PT (Paper 5)

Deliberate PT: For deliberate PT, a similar pattern emerged. Post-hoc tests after the significant ANOVA (F(2.73, 616.99) = 63.95, $p < .001$) showed that the high anthropomorphism condition had again the highest PT ratings, followed by the low anthropomorphism condition (all *ps* < .001). However, in contrast to intuitive PT, no significant difference could be found between the medium-low and medium-high condition ($p = .120$). These results support the uncanny valley effect also exists for deliberate PT ratings, albeit rated generally higher than intuitive PT. A graphical overview on the detected uncanny valley effects is given in Figure 6.2 which shows that both valleys seem to be mostly parallel for the anthropomorphism levels. Further, since PT is a crucial impact factor for robot acceptance and further use intentions, we investigate the impact of anthropomorphism levels, intuitive PT, and deliberate PT on use intention in the following model (Figure 6.3).

6.1.4.3 Mediation Analysis

We used multilevel mediation analyses in a 1-1-1 mediation (Zhang et al., 2009) with the R package mediation (Tingley et al., 2014) and lme4 to check to which degree deliberate and intuitive PT contribute to explain the effects of the robots' anthropomorphism on reuse intention. The intraclass correlation (ICC) indicated that using a multilevel approach is necessary for both deliberate PT (ICC $= .49$) and intuitive PT (ICC $= .15$). The anthropomorphism conditions were dummy coded, with the low condition coded as 0 and the other three robots coded as 1. We used a model with random intercepts and fixed slopes. The overall results of the mediation model are displayed in Figure 6.3.

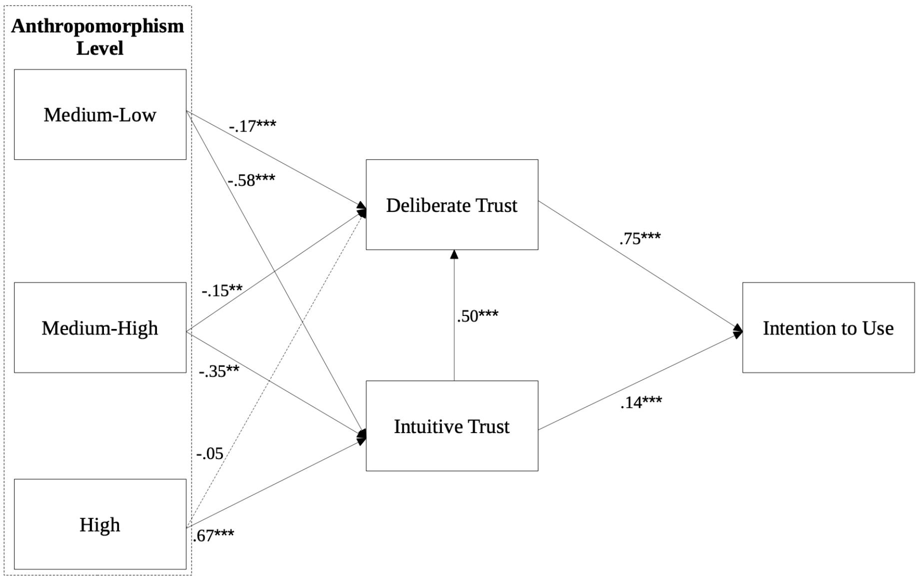


Figure 6.3: Results of the Mediation Model (Paper 5)

Deliberate PT. For deliberate PT, a multilevel model with PT as dependent variable and the anthropomorphism conditions as well as intuitive PT as predictors revealed that the medium-low condition ($beta = -.17$, $p = .001$) and the medium-high condition ($beta = -.15$, $p < .003$) were perceived less trustworthy than the low condition, whereas the high condition did not show a significant difference from the low condition ($beta = -.05$, $p < .32$). Additionally, higher intuitive PT also lead to a higher degree of deliberate PT ($beta = -.50$, $p < .001$). A regression on intention to use showed that both deliberate PT ($beta = .75$, $p < .001$) and intuitive PT ($beta = .14$, $p < .001$)

were positively and significantly related to intention to use, whereas none of the anthropomorphism conditions reached significance (all *ps* $> .19$). After 5000 iterations, the confidence interval for the average causal mediation effects (ACME) did not include zero for the medium-low condition ($beta = -.13$, $CI[-.21; -.05]$, $p = .002$) and the medium-high condition ($beta = -.11$, $CI[-.20; -..04]$, $p < .001$) but did include zero for the high condition ($beta = -.04$, $CI[-.12; .04]$, $p = .32$). Therefore, our results support that increasing anthropomorphism of a robot to a medium level decreases use intention through deliberate PT, whereas there is no direct effect on deliberate PT of increasing the anthropomorphism to a high level.

Intuitive PT. For intuitive PT we proceeded in three steps. In step 1, we checked the anthropomorphism level → intuitive PT → deliberate PT mediation, in step 2 the intuitive PT → deliberate PT → use intentions relationship and finally, in step 3, the anthropomorphism level → intuitive PT → intention to use mediation.

For step 1, the mediation analysis showed a significant mediation effect as well as a confidence interval excluding zero for all three anthropomorphism levels. Specifically, the ACME indicated that seeing the medium-low condition ($beta = -.29, CI[-.37; -.22]$, $p < .001$) and the medium-high condition ($beta = -.18, CI[-.25; -.11]$, $p < .001$) lead to reduced deliberate PT because of reduced intuitive PT for these robots. On the other hand, the high condition increased deliberate PT through intuitive trustworthiness ($beta = .33, CI[.26; .41]$, $p < .001$). For the medium-low condition ($beta = -.17, CI[-.28; -.07]$, $p = .002$) and the medium-high condition ($beta = -.16, CI[-.26; -.05]$, $p = .004$), a direct effect remained, whereas no direct effect could be detected for the high condition ($p = .31$). Therefore, especially medium anthropomorphism levels lying in the uncanny valley seem to negatively mediate PT perceptions.

2. For step 2, the mediation analysis revealed that deliberate PT mediated the positive relationship between intuitive PT and use intentions ($\beta = .38, CI[.34; .42], p < .001$), while maintaining a direct effect ($\beta = .14, CI[.10; .18], p < .001$). Thus, deliberate PT seems to have a higher effect on intention to use, while there is still an effect of intuitive PT on use intentions which is not mediated by deliberate PT.

3. For step 3, there were mediation effects for the medium-low condition ($\beta = -.08, CI[-.11; -.05], p < .001$), the medium-high condition ($\beta = -.05, CI[-.07; -.03], p < .001$), and the high condition ($\beta = .09, CI[.06; .13], p < .001$), while none of the direct effects were significant (all ps $> .190$).

6.1.5 Discussion

The following discussion of our results will be divided into two main aspects. First, we discuss the influence of anthropomorphism on intuitive and deliberate first impression PT and its consequences on further human-robot interactions. Second, we focus on our dual-processing perspective on PT for human-robot interactions and derive a resulting process model of PT evaluations and their impact on use intentions.

6.1.5.1 The Power of Anthropomorphism, or Not?

Despite social robots showing increasingly higher levels of anthropomorphic appearance in recent years, our results indicate that these levels may still be insufficient. That is, although we included highly anthropomorphic robots such as Sophia and Mark1 and participants could see them merely as images (therefore avoiding the possibility that insufficient speech production or behavior reduces anthropomorphism), the uncanny valley was still entered both for intuitive and deliberate PT, and, consequently, use intention. For PT ratings, we could observe that intuitive ratings were consistently lower than deliberate PT ratings, which might point to the reflective system consistently re-processing prior perceptions but positively toward use intentions. This assumption is also supported by our mediation model investigating the mediating effects of intuitive PT on deliberate PT, and both PT variants on use intentions. Whereas both the medium-low and medium-high conditions reduced use intentions through both deliberate and intuitive PT compared to the low anthropomorphism condition, a different picture emerged for the high anthropomorphism condition. Specifically, when comparing the high anthropomorphism condition with the low anthropomorphism condition, we could find no support that deliberate PT was able to explain relevant variance in addition to the intuitive PT. Consequently, our results indicate that deliberate PT entered the "same" uncanny valley as intuitive PT which might be due to intuitive PT itself.

Further following this argument, when considering only the low and high anthropomorphism conditions, an increase in PT could be found for the high anthropomorphism condition. Since this increase is significant for intuitive PT, and further the mediation effect of deliberate PT was not significant for the high condition, the intuitive PT evaluation already seemed to dictate how trustworthiness is to be perceived. Thus, it is assumed that not much re-processing in the deliberate system was necessary. That is, the first impression and thus, the first intuitive evaluation seem to be consistent with the further evaluation for both the obvious, low anthropomorphic robot, and the real human, while this seems not necessarily to be the case for robots imitating humans. A result which challenges our prior assumptions that the intuitive PT evaluation might not be able to directly detect the medium-high robot as robot. An explanation as to why this effect can be observed can be given by neuroscientific studies which show that the neural processing of robots having human or human-like faces requires more cognitive effort than processing real humans or obvious robot faces (Saygin, A.P. et al., 2010). This phenomenon further supports our application of a dual-processing perspective as it reflects (unconscious) decision conflicts and errors within the perception and evaluation process of social robots which might occur primarily in deliberate PT.

When further applying this finding to the design of social robots, it may need to be questioned whether we should design robots like humans. This thought has already been addressed by prior literature, suggesting that the closer the design of a robot gets to a human, the more we constrain the robot's capabilities to those of humans (Duffy, 2003). Furthermore, designing robots like humans also elicits increased social processing and categorizing of these robots similar to human agents, which might lead to severe disturbances in human-robot relationships and further result in increased decision conflicts within the human brain. Consequently, it can be argued that designing robots more machine-like, but still in a way that they are perceived as trustworthy seems to be more reasonable than the aim for high anthropomorphism and creating robots indistinguishable from humans. This thought is further supported by literature stating that robots should be designed according to the tasks they will fulfill and the context in which they will interact with humans (Hayashi et al., 2010; Lohse et al., 2008). Consequently, it might be reasonable to establish design guidelines focusing on social robots within a specific role each – for instance, social robots applied in elderly care might need to meet different requirements than robots which act as language teachers (Belpaeme et al., 2018).

6.1.5.2 The Dual-Processing Perspective on Perceived Trustworthiness

The application of a dual-processing perspective of first impression PT has given us several insights into the formation and influence of PT evaluations on further use intentions. During our mediation analysis, we have shown that intuitive PT acts as mediator between the perceived anthropomorphism level of the robot and the formed deliberate PT. Further, deliberate PT has significant mediating effects between first intuitive PT evaluations and use intentions. Nevertheless, it has been shown that a part of use intentions is also explainable with intuitive PT only. As a result, both evaluation systems might be crucial to consider when conclusions about robot acceptance and use intentions are to be drawn. These use intentions further significantly impact whether humans will actually be interacting with the robot (again) or not. The described procedure represents the main finding of our paper regarding a dual-process approach of trustworthiness perceptions in human-robot interactions and is depicted in the following Figure 6.4.

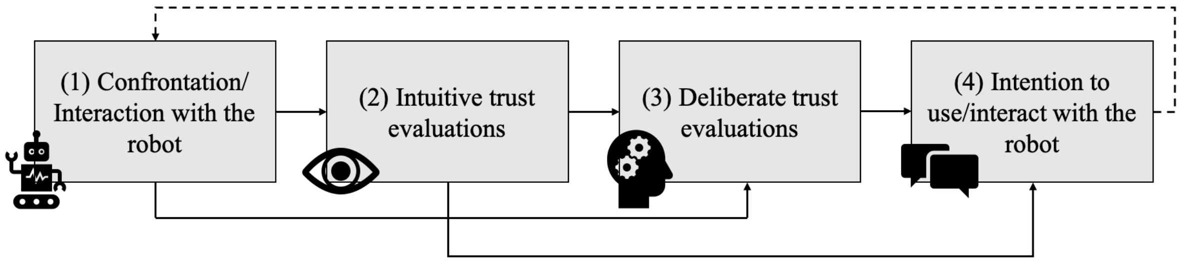


Figure 6.4: A Dual-Processing Model of PT (Paper 5)

Therefore, applying the dual process theory of social reasoning to PT evaluations of robots allows to receive deeper insights into PT formation; and how these mediate the process from the perceived anthropomorphic level of the robot to the use or interaction intentions. For instance, we detected that the impact of anthropomorphism in the real human condition on intuitive PT is especially high and thus, we assume that there is not much re-processing in the deliberate system necessary. In line with this, we showed that deliberate PT seems to more severely impact use intention than intuitive PT. While only the first impression and appearance of robots was tested in this work, this aspect might increase in meaningfulness if actual interaction is investigated. For instance, in case the actual behavior of the robot would be worse than expected, evaluations in the deliberate system might significantly decrease which will further decrease the intention to use and interact with the robot.

6.1.6 Conclusion

6.1.6.1 Summary and Main Findings

In this research work, we have taken a dual-processing theory of PT and investigated its uncanny valley and mediating effects in first impressions of social robots. Since some robots are already designed close to humans, we have focused mostly on robots with high anthropomorphism which are already operating in practice. To gain first insights into how anthropomorphism influences PT ratings, we have focused on the robots' appearance only as a first crucial perception humans get of social robots before interacting with them. Our results show that designing robots in the image of humans does not (or not yet) seem reasonable and that their design should be much more focused and concerned with the tasks the robot will fulfill in society. This might further require specific design guidelines or requirements to be satisfied which should be matched to the role and tasks of the robot.

That being said, the application of a dual-processing perspective on PT has shown us that intuitive PT has significant mediating effects between anthropomorphism level of robots, which were significantly decreased for the human-looking robots than for the machine-like robot. Further, deliberate PT has shown to have the main mediating effect on use intentions, although a part of use intentions are also explainable by intuitive PT alone. To visualize this, we have derived an abstract process model which may help future studies in this area to consider both processes of PT in human-robot interactions.

6.1.6.2 Limitations and Future Work

As every study, this research work does not come without limitations. As its main weakness, it needs to be stated that only first impressions about the visual appearance of robots were investigated in relation to PT. PT itself is, of course, a much more complex construct which might ultimately alter in case of actual human-robot interaction. Since the aim of our study was to make a first step towards investigating the uncanny valley from a dual-processing perspective, we have derived a first model. Future research could therefore investigate this in actual human-robot interactions to further validate our results. In line with this, and as already pointed out in our discussion, it might be reasonable to conduct neuroscientific or NeuroIS studies in this field to receive neural activity as further data input. The application of neuroimaging methods could also help to overcome another limitation of this paper. That is, we distinguished intuitive and deliberate PT mainly by giving time constraints and delays, and by constructing the questions for deliberate PT more complex. While this is proposed as an appropriate method to trigger the two reasoning systems in other studies (Betsch & Kunz, 2008; Glöckner & Witteman, 2009; Sarmany-Schuller, 2010), it cannot be ensured completely that we actually triggered one of the systems at a time. Therefore, by applying neuroimaging methods, further insights into the processing in the human brain can be gained which could help to overcome this potential weakness. Moreover, investigating the neural processing and potential decision conflicts related to the uncanny valley might provide further insights into humans' unconscious perceptions of social robots.

Finally, in our study we only included four different levels of anthropomorphism and female robots (except for the low anthropomorphism condition). Future work might therefore include both male and female robots, as well as more levels of anthropomorphism which might allow to identify design aspects or criteria which provide "thresholds" for when the uncanny valley is entered and when it is left. This would also provide guidance for developing design guidelines for social robots. Given this work, PT seems to be a major predictor for the uncanny valley and should therefore act as one indicator for defining these thresholds. Therewith, a dual-processing perspective on PT is recommendable to receive deeper insights into human-robot interactions.

6.2 Paper 6: Towards Dual Processing of Social Robots

Title	**Towards Dual Processing of Social Robots: Differences in the Automatic and Reflective System**
Authors	Katharina Jahn katharina.jahn@uni-siegen.de Anika Nissen anika.nissen@uni-due.de
Publication Type	Conference Proceedings
Publication Outlet	International Conference of Information Systems
Outlet Information	JOURQUAL3: A
Status	published
Full Citation	Jahn, K., & Nissen, A. (2020). Towards dual processing of social robots: Differences in the automatic and reflective system. *Proceedings of the 41st International Conference on Information Systems (Virtual ICIS), Hyderabad, India*

Table 6.4: Fact Sheet Paper 6

Towards Dual Processing of Social Robots: Differences in the Automatic and Reflective System

Abstract

Social robots increasingly diffuse into our lives in work, health, and private life. However, theoretical approaches that explain how social robots should be designed to maximize experiential and performance-related outcomes in human robot interaction are still rare. To close this research gap, we aim to develop a dual process model of human-robot-interaction with the help of two experiments. Results of the first experiment show that individuals categorize humans and robots differently in the automatic and reflective system, leading to different forms of robotic biases in these systems. With the second experiment, we aim to complement these results from a neurophysiological perspective to gain more insights into cognitive processes during classification and evaluation of robots.

6.2.1 Introduction

The presence and importance of social robots will drastically increase in the future, which might lead to people developing social bonds with robots. Even now social robots can help (elderly) people to go grocery shopping or clean their homes, help with psychological therapy or act as personnel at stores (Mathur et al., 2020; Stock et al., 2019). One important role for the acceptance of social robots is the degree of anthropomorphism they trigger. Anthropomorphism is defined as the degree to which individuals attribute "human-like characteristics, motivations, intentions, and emotions" to "non-human agents" (p. 864 Epley et al., 2007). Cues through which the level of anthropomorphism is perceived and evaluated can be visual, auditory, and behavioral (including cognitive and emotional capabilities) (Pfeuffer et al., 2019). Consequently, as robots are increasingly designed with anthropomorphic features, it stands to question if and how people differentiate between humans and robots and which effects this differentiation might have. For instance, although robots are by now *able* to express emotions, human neural responses to robotic emotional expressions are severely reduced when compared to human emotional expression (Chaminade et al., 2010). This might not only lower the chances of social bonding with robots, but may, even worse, lead to the opposite effect of what is planned to be achieved.

Challenges with the design of social robots become evident in what is called the uncanny valley effect, which proposes that robots within a certain range between high and low anthropomorphic design elicit increased negative reactions to these robots (Mori, 1970; Seymour et al., 2017). This might be due to the prior described reduced emotional experience or negative emotional experience with robots and further, due to grouping and categorization processes between humans and robots. Although research focusing on when and how this effect appears is still ambiguous, two main aspects seem to be identified so far (Kätsyri et al., 2015; MacDorman & Chattopadhyay, 2016; Strait et al., 2017). The uncanny valley might appear a) when inconsistencies between anthropomorphic robots and humans are perceived or b) due to high uncertainty regarding how to place the robot in the social categorization system; a process which is made by humans mostly unconsciously and thus, automatically when they are confronted with another agent. Even though these insights into the uncanny valley effect exist, there is still a research gap on how and when these negative emotions and reactions arise when interacting with social robots. Previous research on the uncanny valley effect has oftentimes investigated perceptions of robots in relation to participants who make an active statement about the group membership (i.e., humanness, prototypicality) of the robot (e.g., MacDorman & Chattopadhyay, 2016, 2017; MacDorman et al., 2009; Seyama & Nagayama, 2007). This, however, has the disadvantage that individuals are directly asked about the group membership or perceptions of the robot. As a result, automatic processes in the distinction between humans and robots might be overseen. In accordance with the *dual process theory* this may become especially problematic, as social cognitive processes have proven to be distinguishable between (i) automatic processes, and (ii) reflective processes (Lobato et al., 2013). Automatic processes happen on the basis of emotion and affect and influence cognition or behavior that occur rapidly on the basis of previous associations and without the need for high cognitive resources, whereas reflective processes describe cognition or behavior based on conscious intentions for which higher cognitive resources are required. Consequently, automatic and reflective processes interact to shape behavior to various degrees in different situations, with most decisions being made in the automatic system (Sloman, 1996). This approach has also been successfully applied to human-robot interaction (HRI) contexts (Bockelman Morrow & Fiore, 2012), which shows that it is generally feasible to consider both systems of the dual processing theory. However, only few studies have used behavioral reaction-time or psychophysiological measurements to assess the automatic processes underlying human perception of social robots (Schindler et al., 2017; Strait et al., 2017). Consequently, this research paper addresses this given research problem by adapting the dual-processing theory to social HRI while considering different levels of anthropomorphism.

In line with this, this research in progress tackles the question *"how far do humans distinguish humans, and high and low anthropomorphic robots in social contexts (i) automatically and (ii) reflectively?"*. Therewith this paper contributes to the body of knowledge by further investigating human processing of robots and humans, which may have significant implications for the design of social robots. Furthermore, this paper helps to gain a deeper understanding of automatic processes which underlie general social processing in humans, which might further become evident in follow-up studies which assess neural cortical activity. To answer our research question, the remainder of the paper is structured as follows: First, related literature regarding social categorization processes and the uncanny valley effect is reviewed. This reviewed literature is already categorized according to the dual process theory (Mathur et al., 2020), and consequently, hypotheses are derived for both automatic and reflective system. After that, a behavioral study is conducted in which general effects between the perception of humans and robots are investigated, also with the distinction between automatic and reflective processing. Finally, the results are presented and their further impact, as well as the planned follow-up study are discussed.

6.2.2 Related Literature: Social Categorization & the Uncanny Valley Effect

First of all, it has to be noted that categorization processes by means of assigning group memberships to different entities is an essential part of human functioning. Humans categorize other entities to make sense of the world around them (Pietraszewski et al., 2014). However, as a result of this categorization, specific biases may arise that, consciously or unconsciously, lead to different emotional reactions towards such categorized groups. In line with this, research on robots has already shown that individuals make similar categorizations within the robot group as within real humans, resulting in stronger biases against these groups, for example relating to race (Eyssel & Kuchenbrandt, 2012). In addition, there is plenty of research showing bias against robots in subjective evaluations of robots or artificially created pictures compared to humans (Gong, 2008).

Thus, previous research has already shown that humans make a clear distinction between entities that merely look almost identical to humans and real humans. In line with this argumentation, we expect that the same automatic and reflective processes apply to robots. Even the most anthropomorphic robots developed up to now can still clearly be distinguished from real humans. Therefore, we expect that overall, robots will be categorized as a separate entity than humans (H1a). In addition, humans are mostly familiar with distinguishing human faces (Pascalis et al., 2002) and assigning more favorable attributes to highly human-looking entities compared to robotic looking entities (Martini et al., 2016). Consequently, we expect that such a (automatic) bias against robots can also be found when comparing robots and humans (H1b). However, biases might not only occur due to automatic processes, but also (and probably especially) reflectively if humans are asked to evaluate the stimuli. Therefore, we also expect a bias against robots in reflective self-reported scales of trust, warmth, and humanness (H2).

Hypothesis 1. Users will automatically a) categorize between humans and robots, and b) show a memory-bias against robots.

Hypothesis 2. Users will reflectively show a bias against robots in self-reported scales if asked to evaluate robots and humans.

Additionally, the prior mentioned uncanny valley effect (Mori, 1970) is one of the most relevant approaches to explain how individuals differentiate between humans and robots. This effect proposes that anthropomorphism helps to increase perceptions of non-human entities up to a certain degree (Mathur & Reichling, 2009). However, if anthropomorphism reaches a certain level between high and low anthropomorphism, the *uncanny valley* is entered (Mathur et al., 2020), which results in more negative reactions against this entity. Only when anthropomorphism is high and becomes almost indistinguishable from a human, the impression becomes more positive again. As we are not interested in investigating the boundaries of when and how the uncanny valley effect appears and can be explained, we use it to derive hypotheses regarding biases humans have against robots. That is, it was shown that even when both entities do not lie within the uncanny valley, less anthropomorphic robots are still rated as least favorable than highly anthropomorphic robots (Mathur et al., 2020). Consequently, we propose that highly anthropomorphic robots will lead to more favorable attribution than robots with less anthropomorphic features. To test this assumption, we will include two types of robots having (i) low anthropomorphic features and (ii) robots having high anthropomorphic features. Thus, we hypothesize:

Hypothesis 3. Users will automatically a) categorize robots with low anthropomorphism more strongly than robots with high anthropomorphism and b) will show a higher automatic bias for robots with low anthropomorphism than for robots with high anthropomorphism.

Hypothesis 4. Users will reflectively show a higher bias for robots with low anthropomorphism than for robots with high anthropomorphism.

While drawing back to the dual processing of social interactions, the empirical study conducted in this work in progress will also aim to capture both automatic and reflective perceptions of (i) low anthropomorphic robots, (ii) high anthropomorphic robots, in comparison to (iii) humans. Automatic reactions will be captured with the employed Who Said What (WSW) paradigm, in which participants view a group discussion of low and high anthropomorphic robots and subsequently have to remember who said what. The memory errors participants make can then be used as an indicator for automatic categorization processes. On the other hand, reflective reactions will become evident in the self-reported scales for humanness, trust, and warmth. The procedure is further described in the following section.

6.2.3 Method

6.2.3.1 Participants, Study Design & Materials

We used two between-subjects conditions: high robotic anthropomorphism and low robotic anthropomorphism. In each of these conditions, humanness was varied by either displaying a human (human condition) or a robot with the respective anthropomorphism (robot condition). Therefore, we used a mixed design using the WSW paradigm. After the online questionnaire was distributed to 74 participants, 38 participants (22 male, 16 female) successfully completed the survey, resulting in 304 data points for the within subject conditions. Participants had to be at least 18 years old to participate, resulting in a mean age of 30.15 years (SD = 11.36). A power analysis was conducted for H3b (the most constrained hypothesis because of the between-subject factor). Previous research in the WSW paradigm has shown a medium effect size for the interaction of two categories ($\eta_p^2 = .06$ Sesko & Biernat, 2010). Looking at the detection of a within-between interaction effect with this effect size in a mixed model, G*Power revealed that 34 participants are sufficient to achieve a power of .80. We used an online experiment with a lottery giving the opportunity to win 3x25 Euro as incentive.

Measures: We assessed the manipulation check for humanness using four self-developed items (“The interviewee is human/is a human/speaks like a human/has a consciousness” Cronbach’s $\alpha = .78$, ranging from .74 to .92 for individual measurements in the within design). To fit the context of our study, in which both robots and humans had to be rated, we carefully selected the three trust items from McKnight et al. (2002b) (“The interviewee is honest/would keep his commitments/is sincere and genuine.”) that could be used both in relation to humans as well as robots, and added an overall item that is commonly used as one-item measure in research on robots ("The interviewee is trustworthy", e.g. Salem et al. (2015). Cronbach's Alpha supported the reliability of the scale (overall $\alpha = .93$; α for individual measures between .76 and .89). Warmth was measured using five items (e.g., "The interviewee is helpful/sensitive/polite/ generous/humble", $\alpha = .93$, ranging from .77 to .87 for individual measures, Eyssel and Kuchenbrandt (2012)). All items were measured on a 7-point scale (1 = strongly disagree, 7 = strongly agree).

Stimuli selection. In order to get a first approximation, we chose to select inanimate images of robots and humans to explore the effects. Following the curves of the uncanny valley effect, it can be assumed that inanimate agents are still able to trigger the same effect, however less extreme than actual interaction would do. Therefore, we assume that if the proposed effects can be measured for still images, they can also be identified for videos or actual interaction, however with higher significance. Therefore, we selected only stimuli with female individuals for the human and high anthropomorphism conditions for two reasons. First, our aim was to gain high ecological validity of our experiment. As robots with high anthropomorphism used today are predominantly female, there is the need to capture how individuals perceive female robots to gain generalizations of experimental effects to practice. Second, research on negative emotions with relation to robots has mainly investigated the effects of robots that can be categorized as male rather than female (Nomura, 2017; Siegel et al., 2009). Further, level of anthropomorphism was determined by robots declared as high anthropomorphic having human faces (see i.e. Figure 1 (b)) while robots declared as low anthropomorphic having humanoid body shape and characteristics like eyes and mouth, but being obviously not human (see i.e. Figure 1 (a)).

Humanness was a within-subjects factor differentiating between humans and robots. For the human condition, four images of humans were selected from the Chicago Face Database (CFD) (namely CFD-AF-255-209-N, CFD-BF-216-132-N, CFD-WF-209-052-N & CFD-WF-238-023-N). To facilitate a higher level of generalization of the images, we used two images of White participants, one image of an Asian person and one image of a Black person. For the robot condition, images of robots used in practice were used. The robotic anthropomorphism was also a between-subjects condition that further differentiated between robots with high and low levels of anthropomorphism. For the high anthropomorphism condition, we selected images of Bina84, Erica, Mark1, and Sophia to align this condition with the races of the human condition. In the low anthropomorphism condition, we selected Walker, Asimo, Nao, and Pepper.

(a) Asimo as a low anthropomorphic human

(b) Mark1 as a high anthropomorphic robot

(c) A human from the CFD for the human condition

Figure 6.5: Illustration of the Different Conditions (Paper 6)

6.2.3.2 Procedure & Data Analysis

The study was set up similar to previous experiments in the *WSW paradigm* (Pietraszewski et al., 2014). Participants were told that the study was about watching interviews with humans and robots and that we were interested in their impressions on the different interviewees. After we obtained informed consent, participants filled out demographic information and completed the trust and novelty seeking questionnaires. Then, they were informed that they would see interviews of different robots and humans for three topics (favorite color, animal, and music) who were interviewed separately. Then, the images of robots and humans and 24 sentences appeared for 15s each. Sentences were randomized within the three topics and the three topics were presented in randomized order. Additionally, robots and humans were randomly assigned to the sentences. After the interviews were over, we used a 1-minute distracting task applied to the European context (seeing a map of the European Union and reporting as many states as possible), similar to previous research in the WSW paradigm. Afterwards, the surprise recall task followed. Here, sentences were presented in a randomized order and participants had to indicate who said what for every sentence. Images of the robotic and human interviewees were presented below the sentence and were randomized across participants but maintained randomization within a participant. After participants had assigned every sentence with an interviewee, each interviewee was rated with the warmth and trust scales adapted to our context, as well as humanness scales. Finally, participants were thanked and debriefed.

Data Analysis – Categorization & Biases. We used the differentiation between two types of errors to test hypotheses H1. To infer information about *categorizations*, the WSW paradigm relies on the errors participants make when remembering who said what. Because participants watched statements made by eight interviewees, they can misattribute the statement an interviewee made to seven other interviewees if they make an error. To unravel categorizations, it is now important what type of error was made, i.e. which group the wrong interviewee belonged to compared to the correct interviewee. Here, two types of errors can be made: within-group (misattributing a statement of a speaker within the robot group) and between-group (i.e. misattributing a statement of a speaker of the robot group with a speaker of the human group). If participants make more within-group errors than between-group errors, it can be taken as an indicator that participants (unconsciously) categorized between the two groups. In the case of no categorization, no difference between within- and between-group errors would be expected. For the robotic *biases*, we used two sources of information. First, the total error count from the paradigm to identify whether participants had a specific (automatic) bias towards remembering contributions originating from humans or robots more correctly. Second, the questionnaire measurements for trust and warmth to identify reflective perceptions as a source of robotic bias. For all of these three dependent variables, separate 2 (humanness: human vs. robot) x 2 (robotic anthropomorphism: low vs. high) mixed ANOVAs were conducted (see Table 6.5 for an overview on means, standard deviations, and conditions).

6.2.4 Results

Automatic Processing: Categorization & Memory Bias. The mixed ANOVA for categorization revealed a significant main effect for humanness ($F(36, 1) = 7.51$, $p = .009$, $eta^2 = .033$) and an interaction effect for humanness and anthropomorphism ($F(36, 1) = 14.53$, $p < .001$, $eta^2 = .063$). Tukey corrected post-hoc tests on the interaction effect revealed a significant difference only for low anthropomorphism, indicating that participants made more within ($M = 8.11$, $SE = .673$) than between errors ($M = 4.56$, $SE = .673$) in the humanness group ($p < .001$), whereas there was no significant difference between within-group ($M = 7.09$, $SE = .679$) and between-group errors ($M = 7.50$, $SE = .679$) for the high anthropomorphism condition ($p = .252$). Therefore, participants categorized between humans and robots only in the low anthropomorphism group but not in the high anthropomorphism group, supporting H1a for the less anthropomorphic robots, and H3a. Looking at the memory bias, a mixed ANOVA for total errors made revealed a main effect for humanness ($F(36, 1) = 11.48$, $p = .002$, $eta^2 = .013$) and an interaction between humanness and anthropomorphism ($F(36, 1) = 8.59$, $p = .006$, $eta^2 = .010$). Tukey corrected post-hoc tests showed that participants only made more errors for statements made from robots ($M = 7.41$, $SE = .630$) than for statements made from humans ($M = 6.26$, $SE = .630$, $p < .001$) in the condition with low anthropomorphism. In contrast, for participants in the high anthropomorphism condition, means were nearly equal between human ($M = 7.25$, $SE = .632$) and robot ($M = 7.34$, $SE = .632$) errors and no significant difference emerged ($p = 754$). Therefore, this result supports H1b, albeit only for the less anthropomorphic robots and as well as H3b.

Reflective Processing: Humanness, Trust & Warmth. The mixed ANOVA revealed a significant main effect for humanness ($F(36, 1) = 84.31$, $p < .001$, $eta^2 = .625$), showing that humans were rated higher in humanness ($M = 5.75$, $SE = .165$) than robots ($M = 3.19$, $SE = .165$). Therefore, the humanness manipulation was successful. Importantly, this shows that robots, regardless of anthropomorphism were correctly identified as not being human. Looking at the subjective trust ratings, this differentiation between humans and robots is further supported. The mixed ANOVA revealed only a main effect for humanness ($F(36, 1) = 16.07$, $p < .001$, $eta^2 = .138$), reflecting higher trust ratings for humans ($M = 4.72$, $SE = .148$) than for robots ($M = 4.01$, $SE = .148$). Neither the main effect for anthropomorphism nor the interaction effect between humanness and anthropomorphism was significant (all ps $> .275$). For warmth, the mixed ANOVA revealed a main effect for humanness ($F(36, 1) = 13.32$, $p < .001$, $eta^2 = .125$), which reflected higher ratings for humans ($M = 4.49$, $SE = .131$) than for robots ($M = 3.89$, $SE = .131$). No other main or interaction effect reached significance (all ps $> .268$). Therefore, H2 is supported in all three included scales of humanness, trust, and warmth. However, H4 has to be rejected for the reflective bias, as no significant interaction effect for anthropomorphism could be found.

	Low Robotic Anthropomorphism		High Robotic Anthropomorphism	
	Robot	Human	Robot	Human
Memory Bias	7.42 (2.41)	6.28 (2.88)	7.35 (3.02)	7.26 (3.10)
Trust	4.12 (1.37)	4.64 (0.74)	3.89 (1.11)	4.80 (1.01)
Warmth	4.00 (1.31)	4.41 (0.61)	3.78 (0.99)	4.56 (0.86)

Table 6.5: Means and Standard Deviations (Paper 6)

6.2.5 Discussion & Follow-Up Study

In this study, we have shown that there is a significant difference in the automatic and reflective processing of robots depending on the degree to which they induce anthropomorphism. We now further discuss our findings in relation to current theoretical approaches below.

Looking at the *automatic* system of human and robot perceptions, our results show that humans categorize robots with low and high anthropomorphism differently. Specifically, our results supported H1 which proposed that the automatic differentiation between human and robots occurs only for robots which are characterized by low anthropomorphism. Highly anthropomorphic robots, on the other hand, were not categorized by the automatic system of participants. This becomes further evident in the interaction effect and thus, H3 was also supported. Interestingly, and maybe due to us using only images as stimuli material, there was no significant automatic distinction between humans and high anthropomorphic robots. Consequently, this might be due to automatic mental state and theory of mind attribution to such robots, even though being aware that the shown stimulus is no human entity. Yet, given the fact that we could not find significant differences in the automatic system for highly anthropomorphic robots, it seems that the degree of anthropomorphism achievable today is sufficient to significantly enhance social acceptance of robots. Consequently, this level of anthropomorphism also seems high enough to provide the high-rated end of the uncanny valley which is also in accordance with findings made by Mathur et al. (2020). Yet, it needs to be stated that there might be a difference between viewing pictures of these robots and actually interacting with them (Seymour et al., 2017). That is, actual interaction with robots might not be anthropomorphic enough to also pass through the uncanny valley. However, our results show and support that when achieving high anthropomorphism and thus, decreasing the differences between humans and robots, the automatic system will not make a clear differentiation and thus, the tendency of social acceptance may be increased. On the other hand, for the *reflective* system for human perceptions of robots, a different picture emerges. That is, participants clearly distinguished between human and robots, independent of the robot's human-likeliness levels (H2 supported, H4 rejected). Consequently, if people are asked to consciously assess and evaluate a robotic agent, no differentiation due to anthropomorphism is made. This shows that people are aware of robots being non-human, even though they might appear human on first sight due to their high anthropomorphism. Consequently, for low anthropomorphic robots, both the automatic and reflective system are consistent in their evaluations, while for the highly anthropomorphic robots, only the reflective system differentiates between humans and robots. Thus, there is a conflict with the automatic system which does not differentiate between the two agents in such cases. This differentiation might be due to the reflective system realizing, that no human mind can be attributed to the robot; an assumption which the automatic system might have made.

To investigate, whether the identified differences in attribution observed in this preliminary study are actually related to the theory of mind, or whether they might be due to other processes, a follow-up study will employ the same experiment procedure as presented in this paper. Albeit, a neuroimaging method will be additionally applied to capture neural responses to the shown images in the paradigm. Therewith, the automatic responses which participants have when confronted with robots will be captured and can be analyzed on an additional data layer. In line with our results, hypotheses for the follow-up study can already be formulated. To achieve this, we draw on literature on grouping and stereotyping processes using neuroimaging methods within real human groups, which revealed significant increases predominantly in the medial prefrontal cortex (mPFC) (Hehman et al., 2014; Molenberghs & Louis, 2018; Morrison et al., 2012). In relation to the dual processing theory used in this paper, the mPFC is also related to attributing mental states and understanding others (Fogassi et al., 2005), which is why it is generally related to the theory of mind and the automatic social processing (Mukamel et al., 2010). Interestingly, a neural activation pattern in the mPFC could also be identified in studies focusing solely on human perception of robots and human-robot interactions. While some of these studies found increases primarily in the left hemisphere (Miura et al., 2009; Strait et al., 2014), other studies found increases bilaterally (Krach et al., 2008; Wang & Quadflieg, 2014). Furthermore, regarding the increasing anthropomorphism of robots, it has been found that more anthropomorphic robots elicit higher increases in the mPFC than less anthropomorphic robots (Krach et al., 2008; Miura et al., 2009). However, when observing anthropomophism changes for real humans by observing avatars, the mPFC is increased for humans compared to avatars (Riedl et al., 2014). Therefore, it can be argued that both for real humans and for robots, increased anthropomorphism leads to mPFC increases. Consequently, further research is required in this area which also takes the uncanny valley effect into account and investigates whether the activation patterns observable in the mPFC appear in relation to this effect. Therefore, we propose the following two hypotheses. The second of which takes two types of robots into account being (i) robots with low anthropomorphism and (ii) robots with high anthropomorphism, which were also already tested in the present study:

Hypothesis 5. A confrontation with low anthropomorphic robots will lead to increased neural activity in the medial prefrontal cortex.

Hypothesis 6. Higher anthropomorphic robots elicit a higher neural increase in the medial prefrontal cortex than less anthropomorphic robots.

With the results from the follow-up study, the first insights we gained within this paper can be further validated and thus, strengthened. Therefore, we plan to repeat the used WSW paradigm using the same stimuli material as well as using video material of the agents. However, in addition to the behavioral measure of the WSW task and the questionnaire measurements, we will further use functional near-infrared spectroscopy (fNIRS) to test Hypotheses 3 and 4 and conduct the experiment in a laboratory environment in contrast to the online experiment. fNIRS is able to capture cortical hemodynamic responses to stimuli in the human brain and is able to detect neural activity in the hypothesized mPFC. While this work in progress has already shown that there are significant differences between the automatic and reflective system when it comes to being confronted with social robots, we hope to identify further, more subtle differences in the neural activity when humans are confronted with robots. However, due to civic circumstances data collection for the fNIRS experiment could not be started yet. Therefore, it is planned to start with this study as soon the situation allows it.

6.2.6 Conclusion

Our research contributes to current literature by proposing a dual process perspective for human-robot research. As with every research, our study has some limitations. One limitation is that we used only images with text and did not investigate actual interactions. Therefore, and in line with assumptions by Mori (1970), the measured effects regarding the uncanny valley might be lower than it would be in actual interactions. Further, the inconsistent attribution between automatic and reflective system for high and low anthropomorphism might also be non-existent in real interactions due to the differences being too high in terms of gestures of mimics. As a result, more research in this area is necessary to validate our results, using videos with audio, or even more immersive technologies like virtual reality in which actual robot interactions can be simulated. Nevertheless, we believe that our results are still valid, because a) our results can be generalized to text- and picture-based experiences with robotic agents (like i.e. chatbots), and b) according to the uncanny valley, similar results are likely to be obtained if robots movements, mimics and behavior is in accordance to the expectations evoked by the here used images. For instance, when robots are so well designed in the future that they cannot be easily distinguished by visual appearance from real humans anymore they will likely mislead our automatic systems. All in all, our research paves the way for future research on the level of anthropomorphism which is necessary for the automatic system to be able to differentiate robots and humans. The results of our current study show clearly that it is necessary to investigate the dual process theory related to human-robot interaction in more detail. Apart from looking at possible moderators in relation to the dual process theory, such as the role of cognitive load for the categorization in the automatic system, research could also significantly benefit from using immersive virtual reality (VR) technologies to achieve allow for actual interactions while still being in a controllable environment. This would further offer the possibility to pre-design robots or change features and consequently, measure humans' reactions to these changes with comparatively low effort.

6.2.7 Acknowledgements

This paper was funded by the Ministry for Culture and Science of the State North Rhine-Westphalia (Project ID 005-1706-0006) and the German Research Foundation (DFG) (Project ID 416228727 – SFB 1410).

6.3 Paper 7: The Effects of Robotic Embodiment on Intergroup Bias

Title	**The Effects of Robotic Embodiment on Intergroup Bias: An Experiment in Immersive Virtual Reality**
Authors	Katharina Jahn katharina.jahn@uni-siegen.de Bastian Kordyaka bastian.kordyaka@uni-siegen.de
Publication Type	Conference Proceedings
Publication Outlet	European Conference of Information Systems
Outlet Information	JOURQUAL3: B
Status	published
Full Citation	Jahn, K., & Kordyaka, B. (2019). The effects of robotic embodiment on intergroup bias: An experiment in immersive virtual reality. *Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm, Sweden.*

Table 6.6: Fact Sheet Paper 7

The Effects of Robotic Embodiment on Intergroup Bias: An Experiment in Immersive Virtual Reality

Abstract

The dissemination of immersive virtual reality (IVR) with body tracking offers new opportunities for communication across distant places by embodying virtual avatars or actual robots. Thus, when people meet via IVR, different forms of embodiment for humans and artificial intelligence become possible, which may influence the perception of the self, the encoding of information and the evaluation of communication partners. In this research-in-progress paper, we draw upon research on self-presence and social identity theory to investigate how robotic embodiment influences intergroup bias in the context of a user who sees a group discussion of individuals embodied in avatars with different degrees of visual and mental humanness. We hypothesize that less humanness leads to higher misattribution of debaters‘ contributions and more negative evaluations of these debaters. Additionally, we assume that this effect is diminished when the users who watch the group discussion are embodied in an avatar with low humanness.

Keywords— Embodiment, Robots, Artificial Intelligence, Intergroup Bias, Laboratory Experiment, Self-presence.

6.3.1 Introduction

As immersive virtual reality (IVR) becomes increasingly affordable, new opportunities for computer mediated communication (CMC) arise. Immersion, as a technological characteristic, describes the degree to which a technology is inclusive, extensive, surrounding, vivid, matching, and provides a coherent plot (Slater & Wilbur, 1997). Through these means, the technology might strengthen the users' illusion of being in a different place (tele-presence) (Slater & Wilbur, 1997; Witmer & Singer, 1998). With the use of full-body tracking, the matching aspect of immersion can be increased by recording individuals' movements and transferring them to a virtual body. Recent research has shown that users can even adopt a robotic body in a distant place or a virtual body in IVR when their body is tracked (Kishore et al., 2016). Thus, when a camera is placed on the head of a robot, individuals can give a speech in front of an audience or talk with other people over a distance in the robotic body. Likewise, when the movements of all communication partners are recorded and head-mounted displays are used, individuals can meet in IVR, for example to instruct someone on how to enact movements.

When individuals communicate over IVR with body tracking, different combinations of interacting with artificial or actual bodies can occur. For a meeting with individuals from different places without the use of robots, communication partners may all have virtual avatars in IVR. When robots are used, communication partners may all be embodied as a robot (e.g. when robots are controlled over a distance in a place to which humans cannot go) or only some are embodied as robots while the others use their actual bodies (e.g. when a colleague working overseas takes part in a team meeting).

These application scenarios have two implications for CMC. First, from the perspective of the message receivers, they are confronted with the message senders who are embodied in a virtual avatar or robotic body - an experience that can vary fundamentally from real-world experiences as visual appearance can be changed to almost any degree in IVR. Importantly, this experience differs from usual CMC with Desktop computers in the higher immersion that is possible with IVR. Second, from the perspective of the senders, they are embodied in a body that can be different from their actual body to a varying degree, which is also an experience that can hardly be recreated without the use of IVR, except for dreaming (Metzinger, 2013).

On the one hand, individuals who are embodied in an avatar or robot with body tracking can experience so-called full-body ownership illusions. Research on body ownership has shown that the body influences the self and behavior of individuals (Kilteni et al., 2015). On the other hand, individuals who interact with individuals who are embodied in an avatar or robot, may categorize the individuals according to their artificial representations. Research has already shown that group membership influences perception of robots, for example with regard to anthropomorphism (Eyssel & Kuchenbrandt, 2012). Robots who possess a human-like appearance are perceived as having mental states (Martini et al., 2016). Additionally, mental schemas of robots that resemble a human in contrast to a box are associated with higher blood pressure and increased negative emotion (Broadbent et al., 2011) and robots resembling outgroup members are evaluated more negatively than robots representing ingroup members (Eyssel & Kuchenbrandt, 2012) and are more likely to be killed in a shooting game (Bartneck et al., 2018).

However, although these initial findings indicate that robotic bodies can be perceived differently than human bodies and intergroup bias for robots seems to exist, previous research mostly investigated reactions to pictures (Bartneck et al., 2018; Eyssel & Kuchenbrandt, 2012; Martini et al., 2016), leaving it unclear how effects would be in real communication situations. This is especially the case when artificial intelligence and humans communicate in artificial bodies, as previous attempts did either manipulate permanent visual aspects of the robots' design (Bartneck et al., 2018; Martini et al., 2016) or information about the robot (Eyssel & Kuchenbrandt, 2012), but not both aspects in one experiment. Furthermore, research investigating the effects of embodying real humans in a robotic body (Kishore et al., 2016) did not investigate whether robot-related intergroup bias could change according to embodiment. Thus, designers of IVR communication systems cannot decide on a theoretical basis which design decisions are the most favourable in a given communication contexts, especially when multiple individuals are present at the same time . We therefore want to address the following research question:

Research Question. How does robotic and artificial embodiment of senders and receivers influence intergroup bias?

To address our research question, we will conduct a laboratory experiment in IVR using a scenario in which the user watches a discussion group in IVR. For this, we vary both the embodiment of the user and the embodiment of the individuals in the discussion group on the dimensions of visual humanness (human vs. robot) and mental humanness (human vs. artificially intelligent). We assess how accurately the users can attribute statements from the discussion to the correct individuals and how the individuals are evaluated regarding affective and cognitive attributions.

The paper is structured as follows. First, we give an overview on the theoretical background regarding body ownership illusions and social identity theory. In the next section, we explain our research model and hypotheses. Subsequently, we describe the experimental design. To conclude, we discuss the implications our research can have for theory and design.

6.3.2 Theoretical Background

6.3.2.1 Self-Presence

Self-presence is defined as "the illusion that one's virtual representation (e.g. avatar) is indeed oneself, that is, inhabiting the virtual body" (Schultze, 2010, p. 438). This construct is closely related to body ownership illusions, which are not constrained on virtual contexts. One typical way of experimentally inducing body ownership illusions is the rubber hand illusion. To trigger the illusory perception that a rubber hand is an individuals' own hand, the rubber hand is placed in front of the individual in the same position as their real hand (which is invisible to them). The experimenter subsequently strokes both the rubber hand and the real hand at the same time. After some time, participants start to perceive that the rubber hand is part of their own body (Botvinick & Cohen, 1998; Kilteni et al., 2015).

With the use of head-mounted displays and body tracking, individuals can embody a virtual avatar, which enables them to experience so-called full-body ownership illusions (Kilteni et al., 2015) that elicit a high sense of self-presence (Schultze, 2010) to the user and subsequently change a range of attitudinal and behavioral variables, even when visual appearance is not photorealistic (Jo et al., 2017). For example, light-skinned individuals who are embodied in a dark-skinned avatar drum in a larger radius (Kilteni et al., 2013) and show reduced intergroup bias (Hasler et al., 2017). On the other hand, individuals who are embodied in an old avatar can show more favourable attitudes towards the elderly and feel closer to them (Oh et al., 2016). Thus, in the context of being embodied in a virtual body in a communication context, and subsequently experiencing high self-presence, it is likely that the design of the body changes how users perceive themselves and others. In the next section, we draw upon the social identity theory and the common ingroup identity model to understand this process.

6.3.2.2 Social Identity and Re-categorisation

According to the social identity theory (Tajfel & Turner, 1986), individuals strive to establish, improve, and maintain a positive self-image. To achieve this, they categorize themselves to groups they belong to (ingroup, e.g. human beings) and compare the value of this group to other relevant groups they do not belong to (outgroup, e.g. animals). As a consequence, to achieve a positive distinctiveness between the ingroup and the outgroup, individuals may engage in discriminatory cognition or behavior towards outgroups and favour the own group over other groups (intergroup bias). However, the categorisations of individuals are not stable and rely on context-specific factors (Ellemers et al., 1999). Therefore, the self-concept of an individual can be understood as a fluid entity.

The common ingroup identity model states that re-categorisation in a common, superordinate group can decrease intergroup bias for cognitive, affective, and behavioral outcomes. This model could be supported both in laboratory (Dovidio et al., 1997; Gaertner et al., 1989) and field studies (Nier et al., 2001). An example of how re-categorisation can occur is displayed in Figure 6.6. If individuals encounter individuals of an outgroup (e.g. a light-skinned individual meets a virtual avatar with dark skin in IVR), they can perceive similarities based on physical appearance between themselves and the out-group member (e.g. the skin of the avatar the user embodies is as dark as the skin of the outgroup member) which helps re-categorising them both in a superordinate group, and can then result in more positive cognitive, affective, and behavioral outcomes and thus, less intergroup discrimination and bias. In the context of communicating with virtual avatars, robotic bodies and artificially intelligence can both represent outgroups for human beings. On the basis of this assumption, we develop our specific hypothesis for the effects of robotic embodiment in communication contexts.

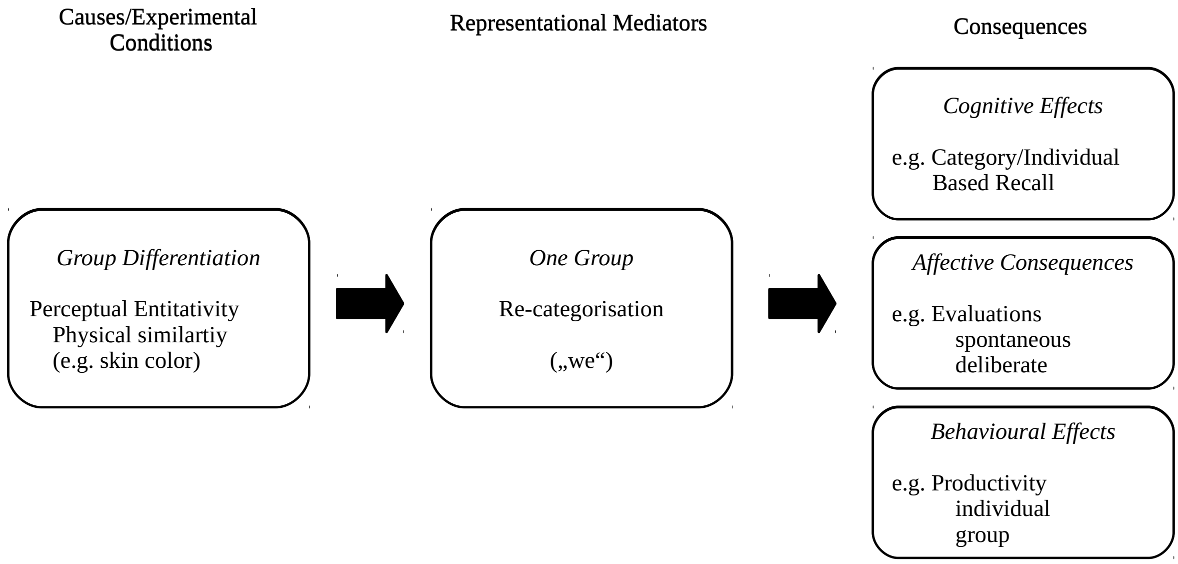


Figure 6.6: Example of the Common Ingroup Identity Model with Perceptual Entitativity as Cause and Re-categorisation as Mediator (adapted from Gaertner et al. (1993) (Paper 7)

6.3.3 Research Model

Our research model is displayed in Figure 6.7. Hypotheses 1-4 are concerned with the effect of *debaters'* embodiment in a group discussion on misattributions and evaluations of the individual who sees the group discussion. Hypotheses 5-8 are concerned with how the embodiment of *the individual who sees the group discussion* interacts with this misattribution and evaluation. Previous research has shown that individuals tend to misattribute contributions of outgroup members in a discussion more frequently than contributions of ingroup members (Sesko & Biernat, 2010; Vescio et al., 2004). As for human individuals, both humans or AI who are embodied in human bodies or AI in generel represent outgroup members, we hypothesize that the same effect will occur for these two groups.

Hypothesis 1. Users will misattribute the contribution of debaters with robotic appearance more frequently than contributions of debaters with human appearance.

Hypothesis 2. Users will misattribute contributions of debaters who possess artificial intelligence more frequently than contributions of debaters with human intelligence.

Besides increased errors in attribution of contribution, we assume that ingroup bias will also show in evaluation of robotic and AI individuals. This is in line with prior research that has shown that individuals tend to ascribe mental states to robotic agents only when they are designed with high humanness (Martini et al., 2016).

Hypothesis 3. Users will evaluate individuals with robotic visual appearance more negatively than individuals with human visual appearance.

Hypothesis 4. Users will evaluate individuals who possess artificial intelligence more negatively than individuals who possess human intelligence.

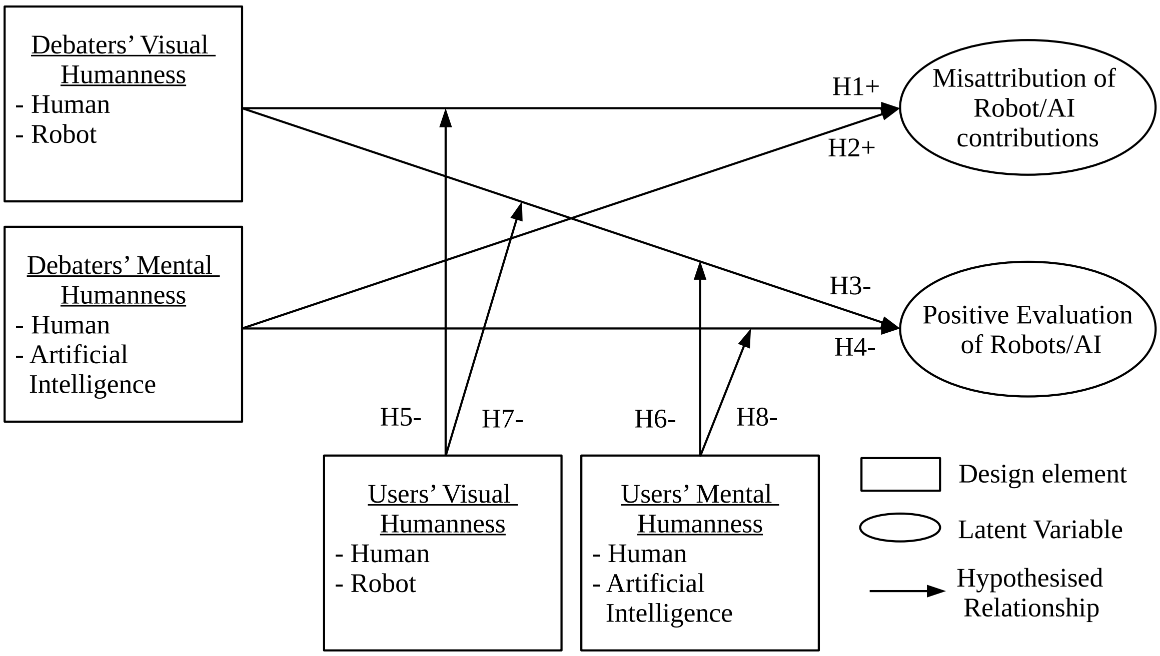


Figure 6.7: Research Model for the Effect of Robotic and Artificial Embodiment on Social Categorisation (Paper 7).

When individuals are embodied in an outgroup avatar, the perception of self-other overlap can change towards the outgroup members (Oh et al., 2016). Additionally, they show increased outgroup mimicry behavior (Hasler et al., 2017). This research indicates that a process of re-categorisation takes place, in which members of the outgroup are perceived as more similar to the self than members of the ingroup, as predicted by the common ingroup identity model (Gaertner et al., 1993). As it is unlikely to assume that individuals will completely change group membership based on their virtual avatar (e.g. from human to robotic), we assume that this re-categorisation process takes place by re-categorisation into a common superordinate group that is more inclusive than the original ingroup (e.g. from the human group towards a group that includes human and robotic individuals). Thus, we hypothesize that misattributions effect will decrease when individuals are embodied in a AI or robotic avatar.

Hypothesis 5. The effect of misattributing contributions of robotic debaters more frequently than human debaters will be diminished when users are embodied in a body with robotic visual appearance.

Hypothesis 6. The effect of misattributing contributions of artificially intelligent debaters more frequently than debaters with human intelligence will be diminished when users are embodied in an avatar with cues for categorising them as artificially intelligent.

Research on body ownership illusions in IVR has shown that embodying the body of an elderly avatar can lead to more favourable attitudes towards the elderly (Oh et al., 2016). Similarly, implicit bias towards dark-skinned people can be reduced when individuals embody an avatar of a dark-skinned person (Peck et al., 2013), and could last even one week after the last experience (Banakou et al., 2016). Thus, in line with H5 and H6, we assume that this categorisation effect is also present in the context of robots.

Hypothesis 7. The negative effect of debaters' visual humanness on evaluation is diminished when users are embodied in an avatar with robotic visual appearance.

Hypothesis 8. The negative effect of debaters' mental humanness on evaluation is diminished when users are embodied in an avatar with cues for categorising them as artificially intelligent.

6.3.4 Method

6.3.4.1 Participants and Design

We will use a fully crossed 2 (users' visual humanness: human vs. robot) x 2 (users mental humanness: human vs. AI) x 2 (debater's visual humanness: human vs. robot) x 2 (debaters' mental humanness: human vs. AI) mixed laboratory experiment to test our research model. Visual and mental humanness of the users' avatar is a between-subjects factor whereas visual and mental humanness of the debaters' avatars is a within-subjects factor. We plan to recruit 130 subjects.

6.3.4.2 Materials

Immersive virtual reality: We will use a HTC Vive head-mounted display with 5 Vive trackers (one for hip, two for feet and hands) and two controllers for interaction with the environment. Participants will be placed in a virtual room with a mirror at one side and a table at the front of the room, enabling participants to watch the panel discussion from the middle of the room.

Figure 6.8: Design of Visual Humanness Condition (Paper 7). A robotic discussion group member is displayed on the left and a human discussion group member is displayed on the right

Users' Avatar, Visual Humanness: We use Adobe Fuse CC (Beta) to create all avatars. The visual appearance of the users' avatar is modelled in both conditions with a human body. However, in the robot condition, the body has a metallic, silvery skin, whereas the body in the human condition has a light-skinned skin tone. We decided to use a human-like appearance out of two reasons. First, previous research indicated that intergroup bias is lowest for human-looking robots compared to box-like avatars. Thus, our results will indicate a minimal difference for intergroup bias for robots. Second, from a methodological point of view, choosing a human-like appearance minimizes visual differences between the conditions. As a result, differences between the conditions can more likely be attributed to the robotic nature of the avatar than to other possible confounders (e.g. other forms of facial expressions when robot has box-like appearance). The gender of the avatar is mapped to the actual gender of the participant to avoid reducing the sense of presence for female participants, which was indicated in previous research (Schwind et al., 2017).

Users' Avatar, Mental Humanness: To implement the mental humanness condition for the users' avatar, the avatar either wears a shirt that is labelled with "AI" (artificial intelligence condition) or "Human" (human condition). Apart from the shirt label, the design of the users' avatar in both conditions is identical. We chose a shirt label because previous research showed that arbitrary, impermanent visual cues, such as cues in clothing, are sufficient to elicit categorisation processes (Kurzban et al., 2001). Thus, this operationalisation allows us to investigate whether these categorisation processes carry over to the user who is embodied in an avatar with this label.

Group discussion Avatar, Visual Humanness: The group membership of the group discussion avatars is created in congruence with the group membership of the users' avatar. Whereas all avatars are modelled after human individuals and have different facial features, avatars in the robotic group possess a metallic, silver skin and avatars in the human condition are light-skinned.

Group discussion Avatar, Mental Humanness: The group membership of the group discussion avatars is created in line with the users' avatar. Thus, avatars in the AI condition wear a shirt which is labelled with "AI" and avatars in the human condition wear a shirt that is labelled with "Human".

Group discussion: To implement the group discussion, we draw upon the "Who said what?" paradigm (Taylor et al., 1978) and use the overall procedure from Sesko and Biernat (2010) but adapt it to IVR. We use a set of 8 virtual avatars with different facial appearance created in Mixamo Fuse (4 robotic, 4 human; 2 of each AI, 2 of each human intelligence) who take part in a group discussion. We develop 16 statements, with each avatar contributing 2 statements to the discussion. The order of the statements is equal for all participants, but the statements that are mapped to avatars vary randomly for each participant.

Evaluation measurements: We use measurements adapted to our context from Eyssel and Kuchenbrandt (2012) to measure warmth and mind attribution with a Likert scale ranging from 1 (not at all) to 7 (very). An overview on the items is given in Table 6.7.

Misattribution measurement: The number of misattributions is calculated as means with 16 possible errors per participant, differentiating between four types of errors:

a) within mental humanness/within visual humanness errors, b) between mental humanness/within visual humanness errors, c) within mental humanness/between visual humanness errors, d) between mental humanness/between visual humanness errors. For example, when a participant makes an error within mental humanness and within visual humanness, it could be that they have misattributed the contribution of an artificial intelligence with robotic appearance to the other artificial intelligence with robotic appearance. On the other hand, an individual that makes an error between mental humanness and between visual humanness could misattribute the contribution of a robot with robotic appearance to another human with human appearance. As for within/within errors, misattribution can only occur for one target, whereas for the other errors, two targets exist, within/within errors are multiplied by two (see, e.g. Sesko & Biernat, 2010; Taylor et al., 1978).

Construct	Item
Warmth	Please rate how you perceive the person regarding the following traits. - helpful - sensitive - polite - generous - humble
Mind attribution	To what extent is this person capable of feeling hungry/joy/pain/fear? To what extent is this person capable of hoping for things? How likely is it that this person has a personality? To what extent is the person capable of being aware of things? How likely is it that this person has a soul?

Table 6.7: Measurement Items Adapted from Eyssel and Kuchenbrandt (2012)

6.3.4.3 Procedure

When participants enter the laboratory, they first answer questions for sociodemographic variables. To ensure that participants understand the meaning of the label that indicates mental humanness, we tell them that our research is the development of AI that can take part in group discussions. For this purpose, we would like them to watch the group discussion of AI and Human participants which use different avatars, but can be distinguished by their label on their shirt.

After the participants put on the trackers and the head-mounted display, they see a room with a mirror and are asked by the experimenter to look around and describe what they see. Next, they are asked to wait in the middle of the room to watch the group discussion. The eight virtual avatars enter the room and go to the table where the panel discussion takes place. Each of the eight avatars then present two statements regarding the topic of why they have lost as a team in an esports game. The statements are presented in random order and are randomized across debaters for every participant.

After the discussion is over, participants are consecutively presented with 32 statements (of which only half were actually present in the discussion) for which they have to indicate whether the statement was present in the discussion and if yes, which avatar had contributed the statement. Finally, they complete measures for warmth and mind attribution for each of the eight debaters and are thanked and debriefed.

6.3.5 Discussion

With the proposed experiment, we expect that we can draw conclusions on the effects that different forms of embodiment in IVR may have. On this basis, technology designers are able to estimate how design decisions regarding robotic or human AI design may influence the users' perception in social interaction. This will be especially relevant when humans are embodied in a robotic body in another place, when people meet virtually or when they need to interact with artificial intelligence. On this basis, future research can investigate related questions, such as the effects of a box-like robotic appearance compared to human-like appearance or the effects of different levels of avatar realism. Furthermore, future research can investigate how intergroup bias is influenced when individuals do not merely watch a group discussion, but also interact within it. We are currently at the stage of programming the group discussion task, animating the characters and synchronising the lip movements to the spoken statements. When this part is completed, we will conduct the experiment and analyse the data.

6.4 Paper 8: Designing Self-Presence in Immersive Virtual Reality to Improve Cognitive Performance

Title	**Designing Self-Presence in Immersive Virtual Reality to Improve Cognitive Performance – A Research Proposal**
Authors	Katharina Jahn katharina.jahn@uni-siegen.de Bastian Kordyaka bastian.kordyaka@uni-siegen.de Caroline Reßing caroline.reßing@uni-siegen.de Kristina Roeding kristina.roeding@uni-siegen.de Bjoern Niehaves bjoern.niehaves@uni-siegen.de
Publication Type	Conference Proceedings
Publication Outlet	NeuroIS Retreat
Outlet Information	–
Status	published
Full Citation	Jahn, K., Kordyaka, B., Ressing, C., Roeding, K., & Niehaves, B. (2019). Designing self-presence in immersive virtual reality to improve cognitive performance - A research proposal. *Proceedings of the NeuroIS Retreat 2019, Vienna, Austria*

Table 6.8: Fact Sheet Paper 8

Designing Self-Presence in Immersive Virtual Reality to Improve Cognitive Performance – A Research Proposal

Abstract

With the increasing availability of immersive virtual reality (IVR) technologies, new opportunities to change individuals' behavior become possible. Notably, recent research showed that by creating a full-body ownership illusion of a virtual avatar looking similar to Einstein, users' cognitive performance can be enhanced. However, although research is quite consistent in reporting that visuomotor synchrony in IVR achieved with body tracking suffices to elicit body ownership illusions that change behavior, it is still unclear whether strengthening these visuomotor illusions with additional technological design elements such as visuotactile feedback can contribute to increase behavioral outcomes even more. In this research in progress paper, we aim to conduct a 2 (physical feedback: low vs. high) x 2 (avatar design: normal vs. high intelligence) between-subjects experiment in IVR to test this assumption. In addition to subjective measures, we use heart rate and electrodermal activity to assess the strength of self-presence induced through the illusions.

Keywords— body ownership illusions, heart rate, electrodermal activity, cognitive performance, physical feedback

6.4.1 Introduction

With the ability to present user's visual, auditory, and tactile senses with completely virtual content, Immersive Virtual Reality (IVR) provides new opportunities to represent the self of an individual. IVR describes a set of technologies that, by enclosing the user with head-mounted displays (HMD) or cage systems heightens sensory immersion. Sensory immersion is a characteristic of the technology, which is high when users are separated into a technology from the real world and their real movements are matched to the virtual environment (Slater & Wilbur, 1997). In contrast to this technological viewpoint, the sense of telepresence describes the psychological perception of the individual, the "illusion of being in a distant place" or "being there" (Schultze, 2010, p. 438), which should arise in individuals when technology provides a high degree of sensory immersion.

In IVR, full-body ownership illusions can be created by combining HMDs with full-body tracking, creating a high degree of self-presence (Schultze, 2010). Self-presence relates to the "Illusion [of] inhabiting the virtual body" (Schultze, 2010, p. 438), when interacting with a virtual body in an environment. Self-presence elicited through body ownership illusions arises when the users' real movements are tracked in real-time and then transferred to a virtual body in the IVR. As a result, the movements of the users' virtual body are displayed in synchrony to the users' real body movements (visuomotor synchrony). This synchrony is sufficient for individuals to experience self-presence (Maselli & Slater, 2013). However, when design elements such as visuotactile or visuomotor synchrony are disrupted, self-presence can be diminished (Kokkinara & Slater, 2014).

Self-presence created by full-body ownership illusions offer many opportunities to enhance behavioral and cognitive outcomes in a beneficial way when working alone or interacting with other people. Individuals embodied in a virtual body with dark skin drum differently (Kilteni et al., 2013) and show decreased racial bias and prejudice (Banakou et al., 2016; Hasler et al., 2017) compared to individuals in a virtual body with white skin. Additionally, individuals embodied in the body of Sigmund Freud show different cognitive processing of problems (Osimo et al., 2015). Furthermore, full-body ownership illusions can even change male users' cognitive performance if they are embodied in an avatar that is associated with high intelligence (Banakou et al., 2018).

Whereas a main factor to elicit full-body ownership illusions with sufficient strength seems to be first person perspective, the strength of body ownership illusions is dependent upon multiple factors. This strength can be measured by biophysiological variables, for example through skin conductance response or heart rate in reaction to a threat (Kokkinara & Slater, 2014; Sanchez-Vives et al., 2010; Tieri et al., 2015). However, whether increasing the effectivity of the body ownership illusion through specific design elements can be used to enhance the cognitive or behavioral outcomes induced through a specific avatar design, is still unclear. Therefore, we want to investigate the following research question to contribute to closing this research gap:

Research Question. How can the interaction between users and virtual avatars be designed to increase users' self-presence and cognitive performance in immersive virtual realities?

To answer our research question, we plan to conduct a 2 (physical feedback: low vs. high) x 2 (avatar design: normal vs. high intelligence) between-subjects experiment.

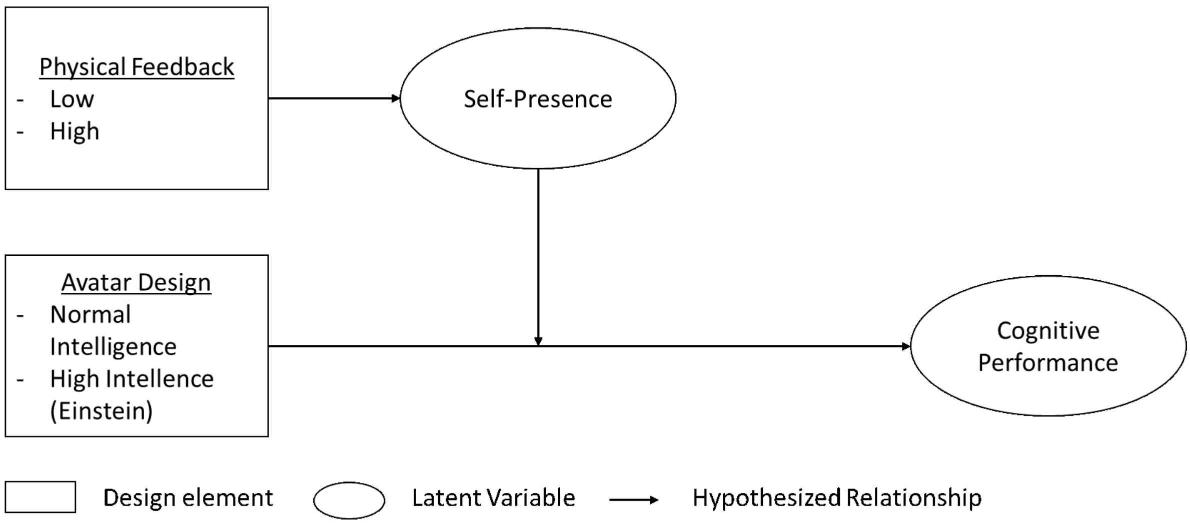


Figure 6.9: Research Model (Paper 8)

6.4.2 Background and Research Model

In this section, we develop our hypotheses based on literature on the antecedents and outcomes of self-presence through full-body ownership illusions. Our research model is displayed in Figure 6.9, which we explain in the following paragraphs.

6.4.2.1 Full-body Ownership Illusions and Effects on the Self

Rooted in the classical rubber hand illusion experiment (Botvinick & Cohen, 1998), in which a rubber hand is touched in synchrony with the individuals' real hand, subsequently arising a sense of ownership over the rubber hand, full-body ownership illusions elicit a sense of ownership over a complete body (Kilteni et al., 2015; Maselli & Slater, 2013). When IVR is used with body tracking, these illusions can create a quite realistic experience of having another body.

From a theoretical point of view, self-presence initiated through body ownership illusions constitutes a passive form of perspective taking (Davis, 1980; Regan & Totten, 1975), in which, rather than imagining to be in the shoes of another person, users can directly experience owning another body (Banakou et al., 2018; Oh et al., 2016). As a consequence, if full-body ownership illusions arise for avatars with specific design elements (e.g. skin color or similarity to a person with competencies in a specific area), individuals cognitive processing and behavior can be influenced (Banakou et al., 2016; Maister et al., 2015). It is assumed that this process occurs by activating resources previously not accessible through this form of perspective-taking (Banakou et al., 2018; Osimo et al., 2015). For example, when individuals were embodied in a virtual avatar of Sigmund Freud when they counselled themselves, they showed more positive mood changes than when they were embodied in a body-scanned version of themselves (Osimo et al., 2015). Additionally, individuals who are embodied in an avatar of Einstein show higher performance outcomes in a cognitive task than when they are embodied in a regular unknown body with which they most likely connect lower intelligence levels than for Einstein (Banakou et al., 2018). We therefore hypothesize that:

Hypothesis 1. Being embodied in a virtual body that is associated with high intelligence leads to higher cognitive performance than being embodied in a virtual body that is associated with normal intelligence.

6.4.2.2 Strength of Self-Presence

Previous research on virtual arms has indicated that self-presence can be induced by synchronous visuomotor stimulation, even when tactile stimulation is absent (Sanchez-Vives et al., 2010). Comparing the effects of visuomotor and visuotactile interaction has shown that the disruption of visuotactile synchrony leads to a lower body ownership illusion (Kokkinara & Slater, 2014). Thus, we suspect that sustaining congruence for visual stimuli coming in contact to the body and touch that is subsequently felt is highly important for keeping the level of self-presence high. This should be especially important in situations in which users have to interact with their hand's multiple times in fine granularity, as this is the case with many virtual reality applications. However, when users' bodies are fully tracked, including their fingers, physical feedback can be incomplete after interaction with virtual objects if no feedback mechanism is implemented in addition to the tracking device. Therefore, we assume that self-presence is higher when physical feedback is presented, and, that, this strengthened self-presence leads to an increased effect of avatar design on cognitive performance.

Hypothesis 2. High physical feedback leads to higher self-presence than low physical feedback.

Hypothesis 3. Self-presence strengthens the effect avatar design has on cognitive performance.

6.4.2.3 Relation of Self-presence to Biophysiological Measures

The level of users' self-reported self-presence seems to be related to biophysiological measures after a threat to the integrity of the virtual body occurs, with the strength of self-presence influencing the strength of the biophysiological reactions to the threat (Maselli & Slater, 2013). Sliding a knife over the artificial body increases electrodermal activity compared to a spoon or asynchronous physical feedback (Petkova & Ehrsson, 2008) and a knife sliding over the body in a condition of first person perspective with synchronous physical feedback results in higher electrodermal activity than third person perspective or asynchronous physical feedback (Petkova et al., 2011). In both studies, these differences were also reflected by the questionnaire items for self-presence. However, other research indicated that synchronous and asynchronous physical feedback is not necessarily reflected by a change in skin conductance response (Kokkinara & Slater, 2014). To gain more insights into these effects, we hypothesize:

Hypothesis 4. Higher levels of self-presence are reflected by an increase in electrodermal activity after the presentation of a threat to the virtual body.

Another biophysiological measure that has been shown to be related to self-presence is heart rate deceleration. After seeing a woman slapping the face of a virtual body from first person perspective, heart rate deceleration increased compared to third person perspective, which was also related to the questionnaire items for self-presence (Slater et al., 2010). Additionally, heart rate deceleration is positively related to self-reported self-presence in a questionnaire after the legs of the virtual body were visually separated (Maselli & Slater, 2013). Thus, we hypothesize:

Hypothesis 5. Higher levels of self-presence are reflected by an increase in heart rate deceleration after the presentation of a threat to the virtual body.

6.4.3 Method

6.4.3.1 Participants and Design

We will recruit 128 male participants to take part in our experiment and use a 2 (physical feedback: low vs. high) x 2 (avatar design: normal vs. high intelligence) between-subjects design to test our hypotheses.

6.4.3.2 Materials and Measures

IVR: A HTC Vive HMD is used to display the virtual environment, which will be designed with Unity 3D. Full-body tracking is implemented with five HTC Vive trackers (2 for hands, 2 for feet, 1 for hip) and hand-tracking is implemented by using Hi5 VR Gloves. Avatars are created using Adobe Fuse.

Electrodermal Activity. We will use electrodermal activity (EDA) as biophysiological measure for self-presence. EDA appear to be most reliably associated with psychopathological states. In line with previous research in the area of body ownership illusions, electrodermal activity is measured in the 6 seconds baseline period and in 2-8 seconds period after the threat (Kokkinara & Slater, 2014). The latency window during which a response will be assumed to be elicited by the stimulus is based on frequency distributions of response latencies to simple stimuli (1-4 sec) (Cacioppo et al., 2007).

Heart Rate. We will use the Polar H7 belt to measure participants' heart rate deceleration. In line with previous research, we measure the mean heart rate for a baseline period of for six seconds before threat presentation and the heart rate after a threat for six seconds after threat presentation (Kokkinara & Slater, 2014; Pollatos et al., 2007). As dependent variable for data analysis, the base measure is then subtracted from the threat measure.

Tower of London Task. This task assesses the level of cognitive performance and is implemented similar to Banakou et al. (2018) in which three differently colored beads on three chopsticks are displayed at descending height. Within three moves, the beads have to manipulate from a predetermined starting position to another set of pins to match the position of the beads in the model. As in Banakou, a point-based algorithm is used to evaluate the experiments (similar to Krikorian et al., 1994.

Questionnaire. We will use the five questions adapted from Banakou et al. (2018) to assess self-presence (body ownership) and agency.

6.4.3.3 Design Elements

Physical feedback. Physical feedback will be designed by providing feedback in form of vibrations through the IVR gloves. Thus, when individuals in the high physical feedback condition touch objects, the gloves vibrate. For individuals in the low physical feedback condition, this vibration is missing.

Avatar design. Avatar design will be operationalized by either using a normal-looking male avatar (normal intelligence condition) or an avatar looking similar to Einstein (high intelligence condition).

6.4.3.4 Procedure

Apart from the physical feedback conditions, the threat to the virtual body, and the psychophysiological measurement, the overall procedure is adapted from Banakou et al. (2018). Participants will be told that they will take part in a study investigating the effects of virtual reality on user experience. Participants will be invited to the laboratory at two time points: during their first visit participants will sign informed consent, and complete measures for self-esteem and cognitive ability and complete the premeasure of the tower of London task. One week later, the IVR session takes place. First, participants are lead into a changing room to put on the HRV belt. Next, the experimenter attaches the electrodes for EDA measurement to the Inside of the medium and Index finger. Afterwards, participants will get instructions on how to put on HTC Vive Trackers and Hi5 VR Gloves. Subsequently, they will put on the HTC Vive HMD and will see a virtual environment which consists of a room with a mirror, a chair, and a virtual body (which either looks like a human or like Einstein, according to the condition) from a first person perspective. When looking in the mirror, participants can see the virtual body mirrored, thus, in a third person perspective. Participants are then asked to get accustomed to the virtual body by moving their body parts and to look around in the virtual room. To engage participants into being in the virtual environment, and to make the physical feedback conditions salient, participants will be asked to complete a task in which they have to locate numbers in the room and sort them in ascending order using their hands. In the high physical feedback condition, participants receive physical feedback when touching the numbers, whereas this feedback is missing for participants in the low physical feedback condition.

In the next part of the experiment, participants will be seated on a chair and asked to answer the virtually presented questionnaires for self-presence (body ownership) and telepresence. After they have finished answering the questionnaire, participants will be told that they have the chance to play a game with a box-shaped robot. In this game, participants are asked to put their right hand on a virtual pad which is tantalized to them by the box-shaped robot. Then, the robot will pull out a knife and starts to stab the knife quickly in the space between the fingers of the participants. This serves as a threat for the virtual body. We chose a game in which the virtual body is not actually hurt because we wanted to refrain from permanently damaging the virtual body, as we expected that this might interfere with the intelligence salience of the Einstein body (participants could remember their experience as threatening rather than as being embodied in the body of an intelligent person). Afterwards, participants take off the HMD and do the post measure of the tower of London task. Finally, participants are thanked and debriefed.

6.4.4 Discussion

With our results, we aim to gain insights into the working mechanisms through which body ownership illusions affect cognitive performance. First, our research contributes to the literature indicating that self-presence in the form of body ownership illusions can be measured by biophysiological variables (Kokkinara & Slater, 2014; Maselli & Slater, 2013) by delivering a more practice-oriented view on physical feedback. Second, we aim to contribute to literature indicating that visuotactile feedback can indeed strengthen self-presence (Kokkinara & Slater, 2014). Third, by testing whether strengthening self-presence can increase, we contribute to practice increasing the knowledge on how immersive virtual reality should be designed to shape behavioral and cognitive outcomes in a beneficial way (Banakou et al., 2018; Osimo et al., 2015; Ott & Freina, 2015).

7 Track 3: Gamification Contexts

7.1 Paper 9: Towards a Unified Theory of Toxic Behavior in Video Games

Title	**Towards a Unified Theory of Toxic Behavior in Video Games**
Authors	Bastian Kordyaka bastian.kordyaka@uni-siegen.de
Authors	Katharina Jahn katharina.jahn@uni-siegen.de Bjoern Niehaves bjoern.niehaves@uni-siegen.de
Publication Type	Journal Publication
Publication Outlet	Internet Research
Outlet Information	Impact Factor: 6.773
Status	published
Full Citation	Kordyaka, B., Jahn, K., & Niehaves, B. (2020). Towards a unified theory of toxic behavior in video games. *Internet Research*, *30*(4), 1081–1102

Table 7.1: Fact Sheet Paper 9

Towards a Unified Theory of Toxic Behavior in Video Games

Abstract

Purpose – Toxic behavior in multiplayer video games diminishes the potential revenue of gaming companies by spreading a bad mood, negatively affecting game play, and subsequently leading to the churn of players. However, research investigating why toxic behavior occurs is still scarce. To address this issue, this study disjunctively tests three different theoretical approaches (social cognitive theory, theory of planned behavior, and online disinhibition effect) to explain toxic behavior and propose a unified theory of toxic behavior.
Design/methodology/approach – In total, 320 respondents participated in a questionnaire study. This study analyzes the data with covariance-based statistics (i.e. regression analysis and structural equation modelling), and the approach is twofold. First, the hypotheses of three theories are disjunctively tested. Second, a unified theory of toxic behavior is proposed.
Findings – The results of this study indicate that online disinhibition best explains toxic behavior, whereby toxic behavior victimization, attitude, and behavioral control also play an important role.
Research limitations/implications – The findings of this study offer an opportunity to better understand a contemporary and especially meaningful form of negative behavior online.
Practical implications – To maintain revenue and popularity, the computer game industry can use the findings of this study to prevent and better address toxic behavior and its negative consequences.
Originality/value – Toxic behavior among video game players is a relatively new and unexplored phenomenon; therefore, this study makes a valuable contribution to the research field by testing the explanatory power of three theoretical approaches and proposing a unified theory of toxic behavior.
Paper type Research paper

Keywords— Psychology, Behavioral sciences, Video games, Toxic behavior

7.1.1 Introduction

Currently, video games are considered the fastest growing leisure market (Chatfield, 2011) and have already become among the most ubiquitous symbols of the contemporary popular culture (Seo et al., 2019). The current estimations suggest that half of the population in Western countries appears to play video games (Muriel & Crawford, 2018). Consequently, the video game industry made more than 137 billion Dollar in 2018 (Newzoo, 2019b), and eSports (i.e. the competitive play of video games) is even considered a university sport with the offering of full-time scholarships (Funk et al., 2018; Hamari & Sjöblom, 2017). Furthermore, gamified approaches have diffused to the working environment in the form of the concepts of serious games (Michael & Chen, 2005) and gamification (Baptista & Oliveira, 2017).

Taken together, video games can be considered a global form of education, entertainment, and sport. To a large extent, the uprising meaningfulness of video games can be attributed to the design element of real-time competition (i.e. multiplayer game modes), which increases players' motivation and enjoyment (Kim & Shute, 2015; Yee et al., 2012). However, the increased use of video games is also associated with new and negative phenomena. One negative phenomenon is toxic behavior (TB), which is an umbrella term used to describe various types of negative behaviors including harassment, flaming, trolling (e.g. gaining enjoyment from intentionally annoying other players), and cheating during games (Adinolf & Turkay, 2018; Kwak et al., 2015; Neto et al., 2017). Since toxic behavior has a direct effect on the churn of players due to a worsened gaming experience and therefore poses a threat to revenue, the video games industry (e.g. Riot Games, Blizzard, and Epic) has attempted to address toxic behavior by teaming up to fight the toxic behaviors of players by establishing the Fair Play Alliance (2018). Unfortunately, current toxic behavior remains a serious problem in various multiplayer online video games.

Surprisingly, research has not been exhaustively addressed toxic behavior. Except for some infrequent attempts (Blackburn & Kwak, 2014; Kwak et al., 2015; Neto et al., 2017), to date, no study has addressed the heart of toxic behavior, which is a theoretical explanation of toxic behavior. To better understand toxic behavior, it is crucial to gain insight into the occurrence of toxic behavior and derive a contemplative type of generalizing thinking about the phenomenon that can be empirically tested. This paper attempts to overcome this shortcoming by proposing the first explanative model of toxic behavior based on relevant knowledge. Accordingly, this paper is guided by the following research question:

Research Question 1. What variables informed by theory explain toxic behavior in multiplayer video games?

To answer our research question (RQ1), we examine the influence of a wide range of potential drivers of toxic behavior in multiplayer video games to provide a holistic point of departure to address toxic behavior. Accordingly, we consult three different approaches that have already illustrated their potential to explain relevant constructs related to the dark side of technology use but have never been used simultaneously in a single study (Kwak et al., 2015; Lapidot-Lefler & Barak, 2012; Pabian & Vandebosch, 2014; Tarafdar et al., 2013; Xiao & Wong, 2013) and test their explanatory power for toxic behavior. Specifically, we use the social cognitive theory (Bandura, 1986) and theory of planned behavior (Ajzen, 1991) as explanations of toxic behavior and propose that the online disinhibition effect (ODE) serves as a mediator (Suler, 2004). We test the theoretical assumptions of all three approaches in form of hypotheses using path modeling based on data from a comprehensive survey and use the results to propose a unified theory. Our procedure is based on the aim to holistically capture the interplay between the psychology of players and the technological environment. We aim to provide a theory-driven explanation of toxic behavior, enabling subsequent research to adequately address toxic behavior and provide a holistic understanding of toxic behavior (e.g. compare the explanatory power of different theoretical approaches and gain insight into spurious relationships). Regarding the context of our study, we use the two most successful and popular MOBA games, namely, League of Legends (LoL) and Defense of the Ancients 2 (DOTA 2), as relevant reflections of the broader context of multiplayer video games. Our objective is to make the following contributions with our study. First, this study should allow academia to better understand toxic behavior, providing a variety of subsequent research opportunities. Second, this study provides practical implications with the opportunity for the video game industry to better address toxic behavior and avoid player turnover while improving the overall game play experience.

7.1.2 Related Work

7.1.2.1 Toxic behavior in multiplayer video games

Consulting pertinent research, toxic behavior is primary addressed in two different contexts. First, a substantial number of previous works examined toxic behavior in organizations as a part of the darker side of organizational leadership (Goldman, 2008; Pelletier, 2010; Reed, 2004). In these studies, toxic behavior is described as an abusive, tyrannical, destructive, and bullying type of leadership with a negative effect on subordinate employees (Green, 2014). Second, research has examined toxic behavior in the context of multiplayer video games (Adinolf & Turkay, 2018; Blackburn & Kwak, 2014; Kwak et al., 2015; Neto et al., 2017). For the purpose of our paper, we focus on the latter context and understand toxic behavior as an umbrella term used to describe different negative behaviors (e.g. harassment, flaming, trolling, and cheating) corroding team effort and harming the game ambiance while playing multiplayer video games (de Mesquita Neto & Becker, 2018; Kwak et al., 2015). We follow the definition provided by Neto et al. (2017), who proposed that toxic behavior is a behavior encountered when a player comes across a negative event during a game that generates anger and frustration, leading to a harmful, contaminated, and disseminated toxic type of communication (Neto et al., 2017). In contrast to the organizational context, toxic behavior in multiplayer video games occurs on the same level of hierarchy. Most toxic behavior is associated with members of the player's own team, and the most common perceived forms of toxicity are flaming and trolling, which occur in almost all games, and, therefore, substantially narrow the gameplay experience (Adinolf & Turkay, 2018).

A key feature related to the occurrence of toxic behavior is the combination of the design elements of team competition and multiplayer exchange, which allows players to attribute failure while playing with other players (Adinolf & Turkay, 2018). Consistently, toxic behavior predominantly emerges over the course of a game as a response to negative events, to discourage existing players (Blackburn & Kwak, 2014) and scare away new players (Shores et al., 2014). In contrast to related and already well researched constructs of the dark side of technology use, such as cyberbullying (Lowry et al., 2019; Lowry et al., 2016), toxic behavior is much more temporary, is a rather normalized part of the ordinary culture of play, and is more anonymous (Kwak et al., 2015). Despite these differences, players may experience psychological and emotional problems, such as anxiety and low self-esteem, resulting from the continuous exposure to toxic behavior because video games occupy a considerable share of their private lives and identities, which allows temporary experiences of toxic behavior to accumulate (Ewoldsen et al., 2012; Kordyaka et al., 2019). In summary, toxic behavior is a serious problem for players and the industry of multiplayer video games.

7.1.2.2 The video game genre Multiplayer Online Battle Arena

One particularly successful genre of multiplayer video games in which the occurrence of toxic behavior can be exquisitely observed is Multiplayer Online Battle Arena (MOBA) games. MOBAs can be considered a subgenre of real-time strategy video games, which is a fusion of longer existing game genres such as action, role-playing, and strategy games. For the purpose of our paper, we refer to the two most currently relevant MOBAs, i.e. League of Legends (LoL) and Defense of the Ancients 2 (DOTA 2). For example, currently, LoL is reasonably considered the most popular video game worldwide (Newzoo, 2019a). This game has more than 115 million active players worldwide (RankedKings, 2019) and generated a revenue of 1.4 billion U.S. dollar in 2018 (Statista, 2019), highlighting the relevance of MOBAs.

At the gameplay level, MOBAs possess different unique characteristics compared to other video games that increase their disposition for toxic behavior. In general, MOBAs are highly dynamic, competitive, and frustrating, and cultivate less autonomy (Johnson et al., 2015). In all games, a player controls a single champion in one of two teams consisting of five players with different abilities. Depending on the outcome of the game, each player wins or loses points in the most frequently played game mode ranked, which are combined to represent their overall level. Thus, collaborating and communication with other players during the game is key to victory. To communicate, players predominantly use text chat and ping commands (i.e. player-relayed alerts that provide gameplay information to the entire team)[1] as communicative sources of toxic behavior.

7.1.2.3 Approaches explaining toxic behavior

To empirically explain toxic behavior, we consult pertinent research that already captured negative behavior in the digital world (see Chapter 7.1.2.1). On the basis of available evidence, we make use of three different theoretical approaches (i.e. online disinhibition effect, social cognitive theory, and theory of planned behavior). Subsequently, we introduce the approaches, define all relevant constructs, and derive the hypotheses.

7.1.2.3.1 Online disinhibition effect.

The online disinhibition effect (ODE) is the perceived lack of restraint an individual feels when communicating online compared to communicating in-person due to decreased behavioral inhibitions (Suler, 2004). Thus, this effect consists of the following two components: benign disinhibition (i.e. positive behavior, such as helping others and showing kindness) and toxic disinhibition (i.e. negative behavior, such as hostile expressions and inappropriate behaviors). The former concept involves the opportunity for individuals to share personal feelings or disclose information that they would be hesitant to share in real life (Lapidot-Lefler & Barak, 2012). The latter concept involves the negative side effect of the loss of inhibition leading to the use of hostile language, swearing, and even threats.

Previous research has used the ODE to explain negative behaviors occurring on the internet (Barlett et al., 2016; Lowry et al., 2019; Lowry et al., 2016). Most such studies found empirical support suggesting that individuals involved in negative behavior exhibit higher levels of disinhibition (Udris, 2014) and that social media use combined with anonymity facilitates negative behavior in digital communities (Lowry et al., 2016), and the mechanisms of moral disengagement have been identified (Runions & Bak, 2015). Interestingly, both concepts (i.e. toxic disinhibition and benign disinhibition) have distinct positive effects on the occurrence of negative behavior (Udris, 2014). We argue that examining the ODE can be a valuable approach because the perception of disinhibition is present in every MOBA game, suggesting that it is likely an especially meaningful predictor of toxic behavior. Specifically, we argue that both forms of disinhibition in the game environment are facilitated by the high degrees of anonymity and invisibility (as a real person) in the game. Additionally, the combination of the design element of competition and the fast-paced nature of multiplayer games leads to a high probability of more automatic behaviors in response to gaming experiences with different valences compared with games without these design aspects. For example, players who experience the negative valence of losing a game or being killed multiple times during a short period of time in the highly competitive gaming environment may cope with this by engaging automatically in insulting others, without feeling the need to control their reactions. Because the players are unlikely to meet each other again (due to the high number of players) or remember their names, social consequences are unlikely, and future toxic behavior is facilitated. Accordingly, disinhibition is a precondition increasing the subsequent occurrence of toxic behavior. Based on the aforementioned discussion, we propose the following two hypotheses related to the ODE:

H ODE.1. Benign disinhibition has a positive effect on toxic behavior.

H ODE.2. Toxic disinhibition has a positive effect on toxic behavior.

Additionally, previous studies have indicated that the ODE is a potential mediator explaining aggressive behavior online. Accordingly, Wu et al. (2017) showed that subjective norms predict toxic disinhibition, and Inocencio-Gray and Mercado (2013) showed that online disinhibition mediates the relationship between perceived behavioral control and environmental factors. We aim to test whether the ODE variables (benign disinhibition and toxic disinhibition) serve as mediators explaining the effect of existing theories (social cognitive theory and theory of planned behavior) on toxic behavior. Due to the innovativeness of this assumption, we do not specify concrete hypotheses and test the influences in an explorative and data-driven fashion. Nonetheless, we propose that psychological variables (related to the SCT and the TPB) influence perceptions regarding technology related variables of the ODE and have distinct direct as well as mediated effects on the dependent variable of our study, toxic behavior.

7.1.2.3.2 Social cognitive theory.

The social cognitive theory (SCT) posits that individuals learn by observing others and explains how this learning impacts behavior (Bandura, 1986). According to this theory, individual learning occurs in a social context and is guided by dynamic and reciprocal interactions between personal and environmental influences on behavior. Personal influences consist of cognitive, affective, and biological factors, and environmental influences consist of social and physical factors occurring in different operationalizations. Central to the SCT is the emphasis on social learning due to external and internal reinforcement (Bandura, 2002). We argue that the SCT is a particularly appropriate approach to explain the occurrence of toxic behavior because behavioral learning in MOBAs largely occurs by observing others. Additionally, toxic behavior is widespread and a rather normalized part of the ordinary culture of play, which increases opportunities to learn corresponding behaviors of toxicity.

Different authors have already used the SCT and identified different constructs linked to the occurrence of negative phenomena on the Internet, which we build on to derive corresponding hypotheses to explain toxic behavior. First, we propose that motivation (i.e. the processes instigating and sustaining goal-directed behavior) is an instrumental driver of toxic behavior in multiplayer games (Xiao & Wong, 2013). This idea is related to previous findings on the meaningfulness of the game environment for players' self-concepts (Ewoldsen et al., 2012; Kordyaka & Hribersek, 2019). Accordingly, players with a high desire for specific motives towards toxicity will try to approve their self-concepts while playing MOBAs and exhibit corresponding behaviors. To identify the most meaningful aspects of motivation in the context of our study, we followed a data-driven approach and carried out a pre-study with 10 players from our university eSports team who were familiar with both games. Based on a list with 16 motivations (Reiss, 2004), we identified power, independence, and status as the most relevant motives regarding toxic behavior. For the purposes of our study, we postulate that players with a high desire for the motives power, independence, and status will be more likely to engage in toxic behavior. Second, we postulate that past victimization experiences of toxicity (i.e. the frequency with which a player has been the target of toxic behavior in the past) support the occurrence of toxic behavior in the future (den Hamer & Konijn, 2015; Fox & Tang, 2014). For this, we refer to the cycle of violence hypothesis, which contends that violent experiences in the past lead to involvement in comparable behaviors in the future (McCord, 1988). Due to the ordinary character of toxic behavior in MOBAs (toxic behavior occurs in almost every game), players frequently have the chance to inflict corresponding victimization experiences. Accordingly, players affected by toxicity are more likely to reproduce toxic behavior in the future. Third, self-efficacy in the game environment is the judgment of a player about his/her capabilities to organize and execute the courses of action required to attain designated types of performance (Xiao & Wong, 2013). Since players in MOBAs are aware of the fact that toxic behavior during games has a detrimental effect on performance, they should aim to avoid the occurrence of toxicity to increase their performance. Accordingly, players who have higher beliefs in their self-efficacy regarding the prevention of toxicity, will refrain from activities that could reduce their performance. Fourth, subjective norms in the game environment describe social influences affecting own behavior (Bandura, 1986). We postulate that the perception of the normative beliefs of important others (e.g. role models and friends who play the same game) regarding toxic behavior influence the assessment of toxicity, leading to higher acceptance and perpetration of toxicity-related behaviors (Bastiaensens et al., 2016). Likewise, in the game context, players will be more likely to engage in toxic behavior if important others approve toxicity. Based on the discussion above, we propose the following hypotheses:

H SCT.1. Motives towards toxicity have a positive effect on toxic behavior.

H SCT.2. Toxic behavior victimization experiences have a positive effect on toxic behavior.

H SCT.3. Self-efficacy towards toxic behavior has a negative effect on toxic behavior.

H SCT.4. Subjective norms approving toxicity have a positive effect on toxic behavior.

7.1.2.3.3 Theory of planned behavior.

The theory of planned behavior (TPB) aims to predict individual's intentions to engage in a behavior of interest at a specific time and context (Ajzen, 1991). As a basis for this prediction, the TPB proposes that the intention to engage in a specific behavior and its subsequent execution can be portrayed as a function of the following three antecedents: the attitude of the individual, which can be understood as an evaluation of the behavior in question; subjective norms, which are the social influences of important others on engaging in the behavior of interest; and perceived behavioral control, which is the difficulty of carrying out the behavior in question. We use the TPB in the context of toxic behavior to provide a structured and well-supported approach with proven usefulness in various contexts. Additionally, we argue that the TPB is a suitable approach to explain toxic behavior because of the references to volitional (behavioral) control regarding the behavior of interest. The TPB has been successfully used in research exploring negative behavior in the online context. Most studies related to this topic examined the educational setting of pupils (Doane et al., 2014; Heirman & Walrave, 2012) and attitude was the strongest predictor of the behavior of interest (Ho et al., 2017). We aim to test the explanatory value of the TPB, because the TPB has been successfully used to predict and explain a wide range of negative behaviors, including smoking, drinking, substance use, etc. Since these theoretical constructs are clearly specified in the relevant TPB literature, we assume that a player's attitude (towards toxic behavior), social factors, and perceived behavioral control regarding the execution of toxic behavior might influence the occurrence of the phenomenon. Specifically, we propose the subsequent two additional hypotheses (for the derivation of the hypothesis concerning subjective norms, see the previous chapter). First, we argue that the attitude (i.e. the positive or negative evaluation) towards toxic behavior is a meaningful predictor of the corresponding group of negative behaviors. Accordingly, a less severe evaluation of toxicity (a more positive attitude) increases the occurrence of toxic behavior. As an example, player behavior in multiplayer games occurs without receiving holistic feedback from others; that is, players who perpetrate toxic behavior do not have a chance to fully understand the negative consequences of their acts. As a result, they underestimate the negative impact of toxic behavior and are more likely to perform related behaviors. Second, behavioral control is the perceived difficulty of performing a behavior of interest and comprises the dimensions of self-efficacy beliefs and controllability (Ajzen, 1991). Accordingly, we postulate that players who perceive toxic behavior as easier to control (have a higher perception of behavioral control) will show lower levels of presentation of toxic behavior because they are aware of the dysfunctional impact of toxicity on performance. Based on this discussion, we propose the following hypotheses related to the TPB:

H TPB.1. See hypothesis SCT.4

H TPB.2. Attitude towards toxicity has a positive effect on toxic behavior.

H TPB.3. Behavioral control regarding toxicity has a negative effect on toxic behavior.

Based on the previously described relationships, we aim to test the following research model and the hypotheses.

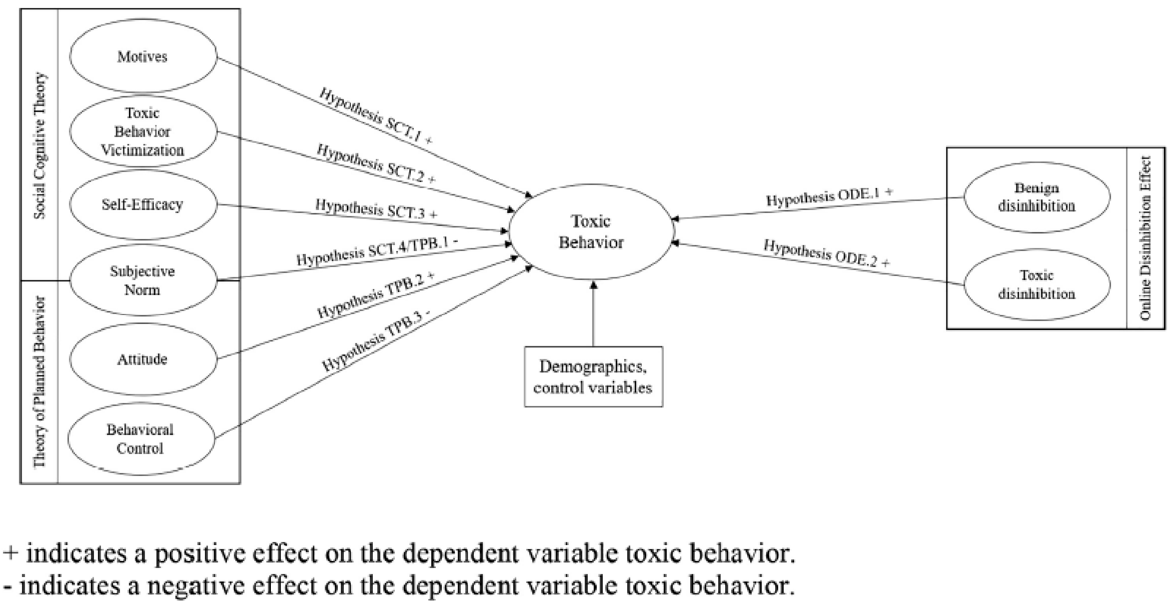


Figure 7.1: Research Model (Paper 9)

7.1.3 Research Methodology

7.1.3.1 Research design

To test our research question (RQ1), we used a cross-sectional design and collected selfreports from players using a digital questionnaire. Subsequently, we analyzed the data with covariance-based statistics (i.e. regression analysis and structural equation modelling) to test the theoretical approaches and explain toxic behavior while controlling for demographics and control variables (see Figure 7.1).

7.1.3.2 Data collection and participants

We conducted a survey of 320 participants using the crowdsourcing marketplace Mechanical Turk (MTurk). To ensure that the participants followed the requirement of playing either LoL or DOTA 2, we asked the participants to specify their three most favorite in-game characters in an open text field. All participants received $1.89 as a reward for participating in our study. The majority of the collected sample were males (214 males, 105 females) and had an average age of approximately 29 years (M = 28.98, SD = 7.24). Most participants were Americans (163), followed by Indians (130), and most participants stated that they had finished their bachelor's degree (80%). Additionally, most participants reported that they used a personal computer as their primary game playing device (59%), followed by video game consoles (27%) and mobile gaming devices (14%). The participants stated that they had been playing video games for 16 years (M = 15.69, SD = 9.23) and play approximately 8 h a week (M = 8.57, SD = 8.98). In addition, 173 participants specified that they predominantly play DOTA 2, while 147 participants predominantly play LoL.

7.1.3.3 Operationalization of constructs

Regarding the operationalization of the constructs, we used empirically validated scales and items adjusted to the context of our study as necessary (see Table 7.2). Most scales used a sevenpoint Likert scale ranging from 1 (“strongly disagree”) to 7 (“strongly agree”). We decided to use the average sum scores of the factors of the relevant items because they can be averaged to reflect the scale, are easy to interpret, and preserve the initial variation in the data (DiStefano et al., 2009). All items used in our study are included in the appendix (see Table 7.5). Additionally, we collected demographic variables (e.g. age, sex, education, and country) and control variables (e.g. hours of play, experience of play, platform, and game) to further prevent unwanted confounding effects on the results.

7.1.4 Data Analysis and Results

7.1.4.1 Validation of the measurement instrument

To validate the measurement instrument, we checked for various validity indicators. First, we checked for face validity by reviewing the measurement instrument using a focus group consisting of three LoL and three DOTA 2 players to check for any ambiguity in the wording or format. All participants in the focus group stated that the wording and format seemed comprehensible. Additionally, we used a sorting exercise (DeVellis, 2016) with two researchers serving as judges. Both researchers were able to correctly place all items on the nine constructs of our study; thus, face validity seemed to be satisfied. Second,we tested our data for common method bias using Harman’s single-factor test (Harman, 1960). No single factor dominated the total variance because the highest eigenvalue explained only 0.31 percent of the variance, indicating that method bias was unlikely to be a concern in this study. Third, we assessed the construct validity of all three theoretical approaches (see chapters 7.1.4.1.1 to 7.1.4.1.3). To assess convergent validity, we referred to existing recommendations in the literature (Gefen et al., 2000) and used the composite reliability (CR) and the average variance extracted (AVE) of the constructs. To test for discriminant validity, we used the Fornell-Larcker criterion, which postulates that a measurement model is supported when the square root of the

Theory / Construct	**Exemplary wording [number of items]**	**Reference**
TTB		
Toxic behavior	"If I get mad during a game, I insult others" [5]	Kordyaka et al. (2019)
ODE		
Benign disinhibition	"I feel like a different person during a game" [7]	Udris (2014)
Toxic disinhibition	"I don't mind writing insulting things about others online because it´s anonymous" [4]	Udris (2014)
TPB		
Subjective norms	"Players who influence my behavior think I should not exhibit TB" [3]	Venkatesh et al. (2003)
Attitude	"disadvantageous – advantageous" [6]	Heirman and Walrave (2012)
Behavioral control	"When I want to prevent myself from exhibiting TB, it is very easy" [4]	de Brujin et al. (2009)
Motives	"Independence" [3(16)][a]	Reiss (2004)
TB victimization	"In the past, other players frequently insulted me" [5]	Kordyaka et al. (2019)
Self-efficacy	"I feel confident understanding the calls to action of other players during a game" [6]	Hsu and Chiu (2004)
Subjective norms	See TPB	

Note(s): We conducted a preliminary study with 10 graduate students to select the three most important of sixteen initial motives (Reiss, 2004) to avoid strain in responding to our questionnaire and content-related inappropriateness (e.g., eating and family)

Table 7.2: Operationalization of the Constructs (Paper 9)

AVE of each construct is greater than the correlations between each construct and the other constructs (Fornell & Larcker, 1981a). Corresponding numbers for all three approaches are presented in Table 7.3. Additionally, we checked for factor loadings and cross-loadings (see Table 7.6). Fourth, we conducted a post study to validate the self-reported level of toxicity by comparing the self-reported and coded data and control our results from social-desirability bias. For this purpose, we used the MTurk worker ID from our last inquiry; we asked 10 participants in our sample to download their last 10 (ranked) games and make them available for us. Due to the rather expensive work assignment, we gave each participant $5.00 as a reward; we further ensured that the selected group of participants reflected the general characteristics of our sample. To ensure the impartiality of the results, we commissioned two students from our eSports group to code the received 100 games as part of their master theses. We introduced the method of qualitative content analysis (Elo & Kyngäs, 2008) and provided information regarding the purpose of our study. Based on text, ping, and behavioral toxicity traces, Coder 1 identified approximately 8 instances of toxicity per game (M = 7.78, SD = 1.33), and Coder 2 identified approximately 6 instances (M = 6.45, SD = 1.14) on average. Based on the total number of participants, we created ordinal ranks of toxicity for the coded and self-reported data and correlated both variables. The results showed a correlation of approximately 90%, which we consider as a clear sign of the validity of the self-reported level of toxicity, consistent with the original literature (Kordyaka et al., 2019).

7.1.4.2 Online disinhibition effect.

To test the measurement model of the ODE, we followed the previously described procedure and inserted the items of the corresponding constructs (i.e. benign disinhibition, toxic disinhibition, and toxic behavior). After inspecting the results of the initial principal component analysis, we excluded the three benign disinhibition items, i.e. BD_2 ("The game is anonymous so it is easier for me to express my true feelings or thoughts"), BD_3 ("It is easier to write things during a game that would be hard to say in real life because you don't see the other's face"), and BD_6 ("I feel like a different person during a game"), and the toxic disinhibition item, i.e. TD_2 ("It is easy to write insulting things online because there are no repercussions"), since these items showed problematic reliabilities and unclear loading patterns. We believe that in the case of the three benign disinhibition items, the results can be attributed to their more complex emotional content compared with that of the other items on the scale. Regarding item TD_2, we assume that the reason for the different correlational structure was most likely that different repercussions exist in the two games of interest, whereas the other three toxic disinhibition items represent factors that are equal across the different games. After the item exclusion, all composite reliabilities exceeded 0.7 (≥0.84), the AVE of each construct was greater than 0.5 (≥0.56), and all items loaded on the intended factors (|>0.65|). All test results met the recommended thresholds, and the convergent validity of the constructs seemed satisfied. Additionally, the square root of the AVE of each construct (≥0.76) was greater than the correlations between each construct and the other constructs (≤0.75), and no meaningful cross-loadings were found (|<0.34|), satisfying the conditions for discriminant validity.

			CR	Mean	SD	1	2	3	4	5
ODE	1	Benign disinhibition	0.84	5.18	1.04	0.75				
	2	Toxic disinhibition	0.84	3.86	1.82	0.26**	0.80			
	3	Toxic behavior	0.93	3.88	1.73	0.19**	0.75***	0.86		
SCT	1	Motives	0.81	4.84	1.30	0.77				
	2	TB victimization	0.88	4.92	1.23	0.46**	0.77			
	3	Self-efficacy	0.88	5.42	1.01	0.30**	0.33**	0.75		
	4	Subjective norms	0.80	5.27	1.30	0.08	0.15**	0.52**	0.81	
	5	Toxic behavior	0.83	3.88	1.73	0.44**	0.41**	-0.04	-0.09	0.87
TPB	1	Attitude	0.85	3.77	1.30	0.85				
	2	Subjective norms	0.82	5.27	1.01	-0.11	0.82			
	3	Behavioral control	0.71	4.69	1.23	-0.27**	0.16**	0.71		
	4	Toxic behavior	0.81	3.88	1.73	0.68***	-0.09	-0.42***	0.81	

Note(s): (a) CR: Composite reliability; (b) Diagonal elements are the square root of the shared variance between the constructs and their measures;
(c) Off-diagonal elements are correlations between constructs.
***p <0.001, **p <0.01

Table 7.3: Descriptive Statistics and Construct Correlations (Paper 9)

7.1.4.2.1 Social cognitive theory.

To test the measurement model of SCT, we inserted the corresponding items of benign disinhibition, toxic disinhibition and toxic behavior. After inspecting the results of the principal component analysis using oblimin rotation, we excluded item SN_2 ("I think players who matter to me would appreciate it if I assist a toxic player") because it was too complex and loaded poorly and relatively equally on more than one factor. We believe that this outcome can be explained by the indirect characteristic of the statement leading to higher levels of ambiguity than the other two items describing subjective norms. After the necessary reduction of the instrument, we conducted another principal component analysis. Accordingly, all composite reliabilities exceeded 0.7 (≥0.80), the AVE of each construct was greater than 0.5 (≥0.56), and all items loaded on the intended factors (j>0.59j). Thus, the convergent validity of the constructs is satisfied. Additionally, the square root of the AVE of each construct (≥0.75) was greater than the correlations between each construct and the other constructs (≤0.52), and no meaningful cross-loadings were detected (j<0.23j), satisfying the conditions required for discriminant validity.

7.1.4.2.2 Theory of planned behavior.

To test the measurement model of the TPB, we used the information derived and the previously described procedure and inserted the remaining items of the corresponding constructs (i.e. attitude, subjective norms, behavioral control, and toxic behavior). All composite reliabilities exceeded 0.7 (≥0.71), the AVE of each construct was greater than 0.5 (≥0.51), and all items loaded on the intended factors (|>0.59<|). All three test results met the recommended thresholds, and the convergent validity of the constructs seemed satisfied. Additionally, the square root of the AVE of each construct (≥0.71) was greater than the correlations between each construct and the other constructs (≤0.68) with no meaningful cross-loadings (|<0.37|), satisfying the conditions required for discriminant validity.

7.1.4.3 Comparison of theories

7.1.4.3.1 Preliminary tests.

To control for any unwanted effects of the demographic and control variables on the dependent variable toxic behavior, we conducted a multiple regression analysis. Therefore, we inserted the demographic (age, sex, education, and country) and control (hours of play, experience of play, platform, and game) variables as predictors of the dependent variable toxic behavior. The regression showed a significant result ($F(8,311) = 11.91$, $p < 0.001$) and explained 22% of the variance (see Table 7.4). In addition, after controlling for the false discovery rate (Benjamini & Hochberg, 1995) using the Bonferroni correction, the variables age (β = -0.27, $p < 0.001$), education (β = 0.26, $p < 0.001$), platform (β = 0.15, p < 0.05), and game (β = 0.17, p < 0.01) had significant effects (all others p $\geq$ 0.24).

7.1.4.3.2 Theory tests.

To disjunctively test the explanatory power of the three theoretical approaches, we used the previously derived information and multiple regression analyses. In the case of the ODE, we used the variables benign disinhibition and toxic disinhibition and the demographic and control variables to explain the dependent variable toxic behavior. The regression equation showed a significant result (F (10,309) = 44.99, p < 0.001) and explained 58% of the variance. After applying the Bonferroni correction, toxic disinhibition (β = 0.68, p < 0.001) and age (β = -0.17, p < 0.01) had significant effects (all others p $\geq$ 0.68). Accordingly, we only found empirical support for one of the two hypotheses related to the ODE (Hypothesis ODE.2: "Toxic disinhibition has a positive effect on toxic behavior").

In the case of the SCT, we used the variables motives, toxic behavior victimization, self-efficacy, and subjective norm and the demographic and control variables to explain the dependent variable toxic behavior. The regression equation had a significant result (F (12,307) = 17.07, p < 0.001) and explained 38% of the variance. After using the Bonferroni correction, motives (β = 0.29, p < 0.01), toxic behavior victimization (β = 0.29, p < 0.01), self-efficacy (β = -0.20, p < 0.01), age (β = -0.21, p < 0.01), and platform (β = 0.15, p < 0.05) had significant effects (all others p $\geq$ 0.34). Regarding our SCT hypotheses, we found empirical indicators confirming Hypotheses SCT.1 (i.e. "Motives towards toxicity have a positive effect on toxic behavior"), SCT.2 (i.e. "Toxic behavior victimization experiences have a positive effect on toxic behavior"), and SCT.3 (i.e. "Self-efficacy towards toxic behavior has a negative effect on toxic behavior"), while only Hypothesis SCT.4 (i.e. "Subjective norms approving toxicity have a positive effect on toxic behavior") did not show the assumed significant relationship. In the case of the TPB, we inserted the variables attitude, subjective norm, and behavioral control and the demographic and control variables as predictors of toxic behavior. The regression equation showed a significant result (F (11,308) = 47.62, p < 0.001) and explained 53% of the variance. After the Bonferroni correction, attitude (β = 0.55, p < 0.01) and behavioral control (β = -0.24, p<0.01) had significant effects (all others p$\geq$0.21). Besides the non-significant effect of Hypothesis TPB.1 (i.e. "Subjective norms approving toxicity increase toxic behavior"), Hypothesis TPB.2 (i.e. "Attitude towards toxicity has a positive effect on toxic behavior") and Hypothesis TPB.3 (i.e. "Behavioral

Predictor variables	Model 1	Model 2 (ODE)	Model 2 (SCT)	Model 2 (TPB)
Age	-0.27***	-0.17**	-0.21**	-0.10
Sex	-0.02	-0.04	-0.03	-0.04
Education	0.26***	0.07	0.93	0.03
Country	-0.11	-0.01	-0.05	-0.01
Hours of play	0.06	0.01	0.05	-0.01
Experience of play	-0.04	-0.01	0.04	0.03
Platform	0.15*	0.07	0.15*	-0.03
Game	0.17**	0.01	0.10	0.10
Benign disinhibition		0.01	0.29**	
Toxic disinhibition		0.68***	0.29**	
Self-efficacy			-0.20**	
Subjective norm			-0.02	0.02
Attitude				0.55**
Behavioral control				-0.24**
R^2	0.23	0.59	0.40	0.54
R^2 adjusted	0.22	0.58	0.38	0.53
AIC	283.00	84.96	212.91	120.77

Note(s): ***p <0.001, **p <0.01, *p <0.05

Table 7.4: Explanatory Power of Theories (Paper 9)

control regarding toxicity has a negative effect on toxic behavior") showed the postulated relationships.

Taken together and based on the quantitative results, the ODE explained the greatest share of variance in toxic behavior (0.58%), followed by the TPB (0.53%), and the SCT (0.38%), while, personal (i.e. motives, toxic behavior victimization, self-efficacy, attitude, and behavioral control) and technological environment (i.e. toxic disinhibition) factors explained toxic behavior. Only subjective norms and benign disinhibition did not reach significance. Examining the Akaike information criterion ("AIC") confirmed the impression that the ODE (AIC = 84.96) showed the most appropriate indicators of fit since it lost less information than the TPB (AIC = 120.77) and SCT (AIC 5 212.91).

7.1.4.4 Proposal of a unified theory of toxic behavior

7.1.4.4.1 Preliminary analysis.

First, we aimed to determine the constructs that significantly explain toxic behavior if we examine all constructs simultaneously. Accordingly, we used a multiple regression analysis using the variables benign and toxic disinhibition, motives, toxic behavior victimization, self-efficacy, subjective norms, attitude, behavioral control, and the demographic and control variables as predictors of toxic behavior. The regression equation showed a significant result (F (16,303) = 35.53, p $<$ 0.001) and explained 63% of the variance. After using the Bonferroni correction, attitude (β = 0.22, p $<$ 0.01), behavioral control (β = -0.15, p $<$ 0.01), and toxic disinhibition (β = 0.42, p $<$ 0.01) had significant effects (all others p $\geq$ 0.27). Second, we searched for predictors of toxic disinhibition. We used a multiple regression analysis including the variables benign disinhibition, motives, toxic behavior victimization, self-efficacy, subjective norms, attitude, and behavioral control as predictors of toxic disinhibition. The regression equation showed a significant result (F (7,312) = 76.13, p $<$0.001) and explained 62% of the variance of toxic disinhibition. After using the Bonferroni correction, toxic behavior victimization (β = 0.16, p $<$ 0.001), attitude (β = 0.59, p $<$ 0.001), and behavioral control (β = -0.15, p $<$ 0.001) had significant effects explaining toxic disinhibition (all others p $\geq$ 0.26). Third, we explored the relationships among the significant predictors of toxic disinhibition. The results showed that toxic behavior victimization and attitude (r = 0.37, p $<$ 0.001) and attitude and behavioral control (r = -0.27, p $<$ 0.001) were significantly correlated (opposed to toxic behavior victimization and behavioral control).

7.1.4.4.2 Unified theory.

To propose a unified theory of toxic behavior, we used the derived information and inserted the data into a structural equation (path) model (Kline, 2015). The results of the path model showed little room for improvement (χ^2(2,320) = 7.95, p = 0.02), which is no longer relied upon as a basis for acceptance or rejection of a model (Schermelleh-Engel et al., 2003; Vandenberg, 2006). Therefore, we assessed additional fit values, which consistently indicated an excellent fit between the theoretical model and empirical model (CFI = 0.99, GFI = 0.99, SRMR = 0.04). Additionally, all predictors accounted for 62% of the variance of toxic behavior. Specifically, attitude (β = 0.26, p$<$0.001), behavioral control (β = -0.18, p $<$ 0.001), and toxic disinhibition (β = 0.50, p $<$ 0.001) predicted toxic behavior. Additionally, toxic behavior victimization (β = 0.21, p $<$ 0.001), attitude (β = 0.63, p $<$ 0.001) and behavioral control (β = -0.15, p $<$ 0.001) predicted toxic disinhibition, while attitude and toxic behavior victimization (r = 0.35, p $<$ 0.001) and

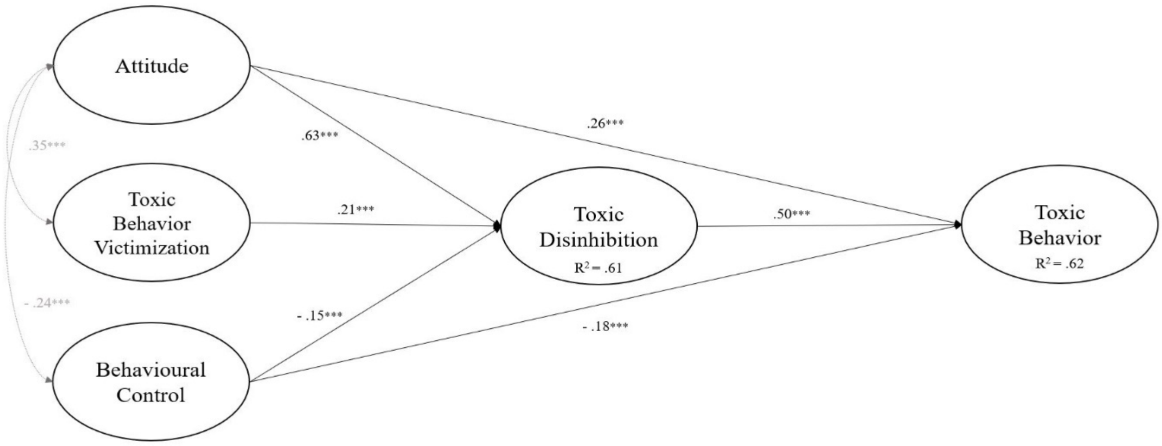


Figure 7.2: Final Path Model (Paper 9)

attitude and behavioral control ($r = -0.24$, $p < 0.001$) showed significant correlations (see Figure 7.2).

7.1.5 Discussion

To answer our research question (RQ1), the present research attempted to integrate three theoretical approaches into a unified theory of toxic behavior. Accordingly, we explored how the factors toxic disinhibition, toxic behavior victimization, attitude, and behavioral control together shape toxic behavior. Having a substantiated type of generalized thinking about toxic behavior that is explanatory and can be tested in empirical research is a meaningful contribution since it allows academia and practice to better understand and address toxic behavior. On this basis, interventions targeting toxic behavior can be developed and specifically designed to change attitudes, experiences of victimization, behavioral control, and toxic disinhibition.

First, the most meaningful predictor of toxic behavior is toxic disinhibition (Hypothesis ODE.2). We consider this finding confirmation of research showing that disinhibition facilitates negative behavior in neighboring contexts (Lowry et al., 2017). However, in contrast to previous findings (Udris, 2014) only toxic disinhibition predicted toxic behavior (Hypothesis ODE.1), which we attribute to an asymmetrical saliency and frequency between positive and negative communication in multiplayer online video games. Additionally, we were able to expand prior findings regarding the mediation effect of the ODE. Therefore, toxic disinhibition not only partially mediated the effects of attitude and behavioral control (Inocencio-Gray & Mercado, 2013) but also fully mediated the effect of toxic behavior victimization. We argue that the feeling of being unidentifiable is particularly relevant for engaging in toxic behavior due to the influence of negative disinhibition.

Second, the SCT variable toxic behavior victimization explains a substantial amount of the variance of toxic disinhibition. Accordingly, the corresponding relationship is positive, indicating that past victimization experiences of toxic behavior lead to higher levels of toxic disinhibition. We interpret this finding to indicate that previous exposure to toxic behavior leads to a higher acceptance (Fox & Tang, 2014). Additionally, in contrast to existing research (den Hamer & Konijn, 2015), toxic behavior victimization only had a fully mediated effect on toxic behavior. We assume that this is the case due to the integration of the TPB and the two direct effects on toxic behavior of attitude and behavioral control, which were not a part of the original literature and might have overshadowed the direct effect of toxic behavior victimization.

Third, the TPB variable attitude is the most meaningful and positive predictor of toxic inhibition. Accordingly, players with a more favorable attitude towards toxic behavior perceive higher levels of toxic disinhibition. Although previous studies have found relationships between the components of the TPB (Inocencio-Gray & Mercado, 2013; Wu et al., 2017) and disinhibition, we interpret our finding as an extension to the context of negative behavior in multiplayer online video games. Additionally, our findings show a direct and positive effect of attitude on toxic behavior (Hypothesis TPB.2). This finding is consistent with previous research indicating that attitude towards the valence of the behavior of interest is a strong predictor of that behavior (Doane et al., 2014).

Fourth, the TPB variable behavioral control is a negative predictor of toxic disinhibition in our unified theory. Werefer to previous research, that has shown that TPB variables have the potential to explain toxic disinhibition (Wu et al., 2017) and, more concretely that a higher perception of behavioral control is associated with lower perceptions of toxic disinhibition. Additionally, we found a direct negative effect of behavioral control on toxic behavior (Hypothesis TPB.3). This finding is consistent with research from the neighboring context, in which the perceived level of anonymity is interpreted as an inverted form of behavioral control during the occurrence of negative behavior in an online context (Kowalski et al., 2014). Potential explanations of toxic behavior might be that due to perceived anonymity, players feel invisible, which removes concerns regarding being caught and socially punished for expressing toxic behavior.

7.1.5.1 Implications for theory

Our study has different theoretical implications. First, by closing the research gap of a missing theoretical explanation of toxic behavior, we contribute to negative online behavior research online representing a contemporary phenomenon. In contrast to research regarding negative online behavior, our study provided empirical indicators suggesting that the integration of three different theoretical approaches (SCT, TPB, and ODE) is promising to adequately and holistically capture toxic behavior. Our results indicate that psychological (i.e. attitude and behavioral control), environmental (i.e. toxic behavior victimization), and technological (i.e. toxic disinhibition) constructs, as well as their interplay explain toxic behavior. Accordingly, three well-tested and established sets of constructs can be applied to the context of toxic behavior in multiplayer online video games using their corresponding quantitative instruments. Future research could build upon the identified relationships and pursue an understanding of toxic behavior.

Second, while our study focuses on a specific video game genre, the results can be generalized to other video game genres, at least to a certain extent. MOBAs represent a particularly successful contemporary game genre and can be considered a stereotype instantiation of a disruptive movement within the industry of video games. Hence, our results can form a baseline for the theoretical description of toxic behavior. Accordingly, we argue that researchers could build upon this study by examining negative behavior in video games in which real time interaction is key to success and toxic behavior emerges as an issue. Additionally, our findings can contribute to existing research concerning negative behavior on the internet, which mainly focuses on the contexts of schools and social media (Doane et al., 2014; Lowry et al., 2016). With our study, we expanded this view by illustrating the interplay among three different theoretical approaches as explanations of a new form of deviant behavior, which could be tested in more general contexts.

Finally, we challenge some findings from previous research in neighboring contexts since we found some unique patterns in our data. For example, the non-significant effect of the subjective norms is a part of both the SCT and TPB (Bastiaensens et al., 2016; Heirman & Walrave, 2012; Wu et al., 2017). To explain this finding, we refer to the unique characteristics of toxic behavior, i.e. toxic behavior occurs in the short terms, is accepted by the community, and does not necessarily require intentions to harm others. Accordingly, we understand toxic behavior as a much more automatic and rather subconscious phenomenon that occurs in stressful situations in competitive gaming. As a result, the influence of significant others is much more internalized and not as conscious during the occurrence of toxic behavior.

7.1.5.2 Implications for practice

Toxic behavior in video games has genuine negative consequences because players who experience aggressive activities may choose to leave the game or initiate more toxicity in return, which may lead to a downward spiral. Accordingly, controlling toxic behavior is critical for the sustained success and survival of video games. Our findings provide various points of reference to better address toxic behavior. First, given the strength of the predictors, toxic disinhibition and attitude, we recommend the use of the concept of associative learning on a level of design to better control impulsivity and poor risk assessment of players. This control could be achieved by underlining the negative consequences of toxic behavior (rewards or punishments) more strongly by providing respected role models who disregard toxic behavior (due to identification) or underlining that the quality of communication with team-mates during a game is directly linked to performance (Neto et al., 2017). Concrete design interventions include giving players the opportunity to play with less toxic players if they accept longer waiting times during the process of match making, the chance to skip such games without losing any points or the establishment of an official code of conduct.

Second, the explored information regarding the fully mediated effect of toxic behavior victimization could be used to better address toxicity. Since toxic behavior predominantly emerges over the course of a game as a response to negative events (Blackburn & Kwak, 2014), the video game industry could provide toxic-specific statistics after each game and offer other players a more holistic feedback system. Another way to address this aspect would be to provide additional chances to participate and examine changes in the game experience throughout each season (i.e. patches).

Third, to increase the level of player behavioral control, cognitive behavioral and self-recognition strategies could be used (Watanabe & Fukuta, 2017). To engage problematic players in related programs and develop anger management tools (e.g. identify the triggers of one's anger, recognize one's warning signs, and change the way one thinks), the game industry could use various starting points. On the one hand, this control could be achieved via positive reinforcements (e.g. providing specific awards to players who do not exhibit toxic behavior). On the other hand, negative punishments (e.g. bans and restrictions) could increase the motivation to participate in such programs. Additionally, the industry could use the technique of co-teaching between players. Concretely, experienced players who know how to address toxic behavior could be used and motivated to educate younger, less experienced players regarding the management of toxic behavior by providing specific incentives.

7.1.6 Limitations and Outlook

Similar to all studies, this study is not without limitations. First, we examined two different multiplayer video games from the MOBA game genre. Thus, future research is needed to determine whether these results can be generalized to other game genres. Additionally, neighboring contexts dealing with negative behavior on the internet can deductively test the (external) validity of our data-driven theory proposal to better understand the phenomena of interest. Second, in response to the rapidly changing environment of video games, there are additional potential factors that could lead to toxic behavior. Future studies need to discover these and test factors in relation to the results of our study. One fruitful avenue would be a more granular examination of the interplay between the online and offline identities of players and how they interact with each other during the occurrence of toxicity. Fourth, our methodological approach identified only correlational relationships. However, future research could build upon and test our findings for causality using experiments. In addition to the post-test of the toxic behavior scale, we mostly used self-reports of players. Accordingly, we encourage future studies to triangulate data from different sources of variance and explore similarities and differences.

Note 1. Although voice chat is possible, it is not available to all five team players in the ranked game mode.

Appendix

Construct		Wording	Reference
TB		"If I get mad during a game, I. . ."	Kordyaka et al. (2019)
Toxic Behavior	TB_1	". . .intentionally interrupt others while they are writing."	
	TB_2	" . . .hold others responsible making own mistakes."	
	TB_3	". . .take away resources belonging to others."	
	TB_4	". . .insult others."	
	TB_5	". . .criticize others."	
ODE		"Please indicate how much you agree with the statements regarding the play of League of Legends/DOTA 2"	

Benign disinhibition	BD_1	"It is easier to connect with others through the game than talking in person"	Udris (2014)
	BD_2	"The game is anonymous so it is easier for me to express my true feelings or thoughts."	
	BD_3	"It is easier to write things during a game that would be hard to say in real life because you don't see the other's face."	
	BD_4	"It is easier to communicate during a game because you can reply anytime you like."	
	BD_5	"I have an image of the other players in my head when I read their messages."	
	BD_6	"I feel like a different person during a game."	
	BD_7	"I feel that during gamesI can communicate on the same level with others who are older or have higher status."	
Toxic disinhibition	TD_1	"I don't mind writing insulting things about others online because it's anonymous"	Udris (2014)
	TD_2	"It is easy to write insulting things online because there are no repercussions."	
	TD_3	"There are no rules online therefore you can do whatever you want."	
	TD_4	"Writing insulting things online is not TB [R]."	
SCT		"Please indicate how much the subsequent motives apply to you as a person while playing the specified game"	Reiss (2004)
Motives	M_1	"Power"	
	M_2	"Independence"	
	M_3	"Status"	

TB victimization		“In the past other players frequently. . .”	Kordyaka et al. (2019)
	V_1	“. . .intentionally interrupted me while I were writing.”	
	V_2	” . . .held me responsible for making own mistakes.”	
	V_3	“. . .took away resources belonging to me.”	
	V_4	“. . .insulted me.”	
	V_5	“. . .criticized me.”	
Self-efficacy		“I feel confident. . .”	Hsu and Chiu (2004)
	SE_1	“. . .completing my tasks while playing a game.”	
	SE_2	“. . .visiting the store for buying items.”	
	SE_3	“. . .navigating over the map using pings during a game.”	
	SE_4	“. . .knowing information about the most recent changes/patches in the game”	
	SE_5	“. . .knowing information about the most recent patches and software changes.”	
	SE_6	“. . .understanding the calls to action of other players during a game.”	
Subjective norms	SN_1	“Most players who are important to me do not perpetrate TB.”	Venkatesh et al. (2003)
	SN_2	“I think players who matter to me would appreciate it if I assist a toxic player [R].”	
	SN_3	“Players who influence my behavior think I should not exhibit TB.”	
TPB Subjective norms		See SCT	
Attitude		I perceive Toxic Behavior as:”	Heirman and Walrave (2012)
	A_1	“bad – good”	
	A_2	“foolish – wise”	
	A_3	“dislike – like”	

	A_4	"unpleasant – pleasant"	
	A_5	"harmful – not harmful"	
	A_6	"disadvantageous – advantageous"	
Behavioral control		"When I want to prevent myself from exhibiting TB. . ."	de Brujin et al. (2009)
	BC_1	". . .it is very easy."	
	BC_2	". . .it is very difficult [R]."	
	BC_3	". . .I am very likely to succeed."	
	BC_4	". . .I am very likely to fail [R]."	

Note(s): [R] describes reverse coded items

Table 7.5: Questionnaire Items

ODE	Benign disinhibition	Toxic disinhibition	Toxic behavior		
BD_1	0.74	-0.04	0.06		
BD_4	0.79	0.20	0.18		
BD_5	0.70	-0.20	-0.13		
BD_7	0.77	-0.20	-0.09		
TD_1	-0.04	-0.65	0.34		
TD_3	0.12	-0.84	0.02		
TD_4	-0.03	-0.80	0.18		
TB_1	0.02	-0.21	0.75		
TB_2	0.02	0.04	0.89		
TB_3	0.02	-0.17	0.77		
TB_4	-0.02	-0.08	0.84		
TB_5	0.01	0.04	0.92		
SCT	Motives Motives	TB victim-ization	Self-efficacy	Subject-ive norms	Toxic behavior
M_1	-0.78	-0.05	-0.05	-0.08	0.09
M_2	-0.82	-0.03	0.05	0.10	-0.11
M_3	-0.71	-0.04	0.12	-0.02	0.18
V_1	-0.23	-0.59	-0.01	-0.07	0.20
V_2	-0.05	-0.74	0.10	0.00	-0.05
V_3	0.05	-0.85	-0.02	0.06	0.01
V_4	0.05	-0.82	0.16	-0.08	-0.03
V_5	-0.05	-0.81	-0.16	0.10	0.08
SE_1	0.11	0.04	0.72	0.14	0.08
SE_2	-0.10	0.01	0.65	0.08	0.00
SE_3	0.08	-0.07	0.73	0.07	-0.02

SE_4	-0.09	-0.10	0.81	-0.12	-0.08
SE_5	-0.15	0.07	0.69	0.05	-0.02
SE_6	0.04	-0.03	0.86	-0.05	0.03
SN_1	0.03	-0.13	0.01	0.82	-0.05
SN_3	-0.07	0.08	0.09	0.81	0.04
TB_1	-0.11	0.02	-0.03	0.00	0.86
TB_2	-0.03	-0.01	-0.03	0.05	0.84
TB_3	-0.13	0.01	-0.02	-0.07	0.84
TB_4	0.08	0.01	0.05	-0.01	0.92
TB_5	0.10	-0.07	0.03	-0.01	0.90

TPB	Attitude Attitude	Subject-ive norms	Behavioral control	Toxic behavior
A_1	0.91	-0.04	0.04	0.05
A_2	0.92	-0.02	-0.01	0.01
A_3	0.85	-0.04	0.02	0.11
A_4	0.93	-0.06	0.04	0.02
A_5	0.86	0.01	-0.03	-0.09
A_6	0.92	-0.02	0.06	-0.01
SN_1	-0.04	0.84	0.06	-0.13
SN_3	-0.12	0.79	0.09	-0.02
BC_1	0.22	0.21	0.79	-0.01
BC_2	-0.23	-0.29	0.60	-0.37
BC_3	0.05	0.17	0.85	0.21
BC_4	-0.31	-0.34	0.59	-0.29
TB_1	0.16	-0.02	-0.01	0.78
TB_2	-0.03	-0.03	0.07	0.89
TB_3	0.15	-0.07	0.02	0.80
TB_4	-0.05	-0.01	-0.02	0.92
TB_5	-0.06	-0.03	0.02	0.93

Table 7.6: Loadings and Cross-loadings of Items

7.2 Paper 10: Individual Boundary Management

Title	**Individual Boundary Management: An Empirical Investigation on Technology-Related Tactics**
Authors	Katharina Jahn katharina.jahn@uni-siegen.de Michael Klesel michael.klesel@uni-siegen.de Kristina Lemmer kristina.lemmer@uni-siegen.de Andreas Weigel andreas.weigel@students.uni-siegen.de Bjoern Niehaves bjoern.niehaves@uni-siegen.de
Publication Type	Conference Proceedings
Publication Outlet	Pacific Asia Conference on Information Systems
Outlet Information	JOURQUAL3: C
Status	published
Full Citation	Jahn, K., Klesel, M., Lemmer, K., Weigel, A., & Niehaves, B. (2016). Individual boundary management: An empirical investigation on technology-related tactics. *Proceedings of the 20th Pacific Asia Conference on Information Systems (PACIS 2016), Chiayi, Taiwan*

Table 7.7: Fact Sheet Paper 10

Individual Boundary Management: An Empirical Investigation on Technology-Related Tactics

Abstract

Elevated through the increasing digitalization, employees are expected to be available always and everywhere. According to boundary theory, individuals can manage their boundaries between work and private life on a continuum of integration and separation. As individuals have different preferences for integration or separation, they are implementing IT tactics to meet their preferences. However, there is a lack of research addressing this topic. Therefore, we used an exploratory approach using tools from grounded theory in order to detect IT-related tactics which employees use to manage their boundaries between work and private life in a way that is in line with their preferences. We identified six tactics that varied in their ability to foster integration or separation and could be administered either manually or automatically. These tactics ranged from physical detachment in which employees separate work and private life manually through creating distance between the device and themselves up to dynamic filters with which the device automatically filters messages from different people and lets only relevant messages come through

Keywords— Boundary Theory, Boundary Management, Individual IT Tactics

7.2.1 Introduction

Due to the technological evolution of mobile technologies including smartphones, tablets and wearables, job-related tasks can be performed nearly anywhere and anytime (Karanasios & Allen, 2014; Reyt & Wiesenfeld, 2015). According to a forecast from the International Data Corporation (IDC) in 2015, mobile worker population will grow steadily in the next years, increasing from ca. 96 million in 2015 to over 100 million mobile workers in 2020 – only in the U.S. By the end of the forecast period, mobile workers will account for almost three quarters of the total U.S. workforce (IDC, 2015). Key drivers behind the growth of mobile workers includes reduced prices of smartphones and tablets combined with the growing acceptance of corporate bring your own device (BYOD) programs in organizations (IDC, 2015). Additionally, technological innovations such as wearables, near-field communications (NFC), voice control and augmented reality are enabling workers to increase their productivity by optimizing communication along organizational workflows (IDC, 2015).

Based on the technological advancement, there is a fundamental change with regard to workplace design, i.e. working times are getting more flexible and workplaces are getting location-independent. Therefore, organizations are facing new demands, norms and a cultural change. Concepts like BYOD ("Bring Your Own Device") and IT-Consumerization (Koeffer et al., 2015) are well-known examples and force organizations to rethink their policies and cultures with regard to the organizational use of technology.

Previous research on the use of mobile technologies has found both positive and negative effects on an individual's work and private life domain (Allen et al., 2014). Besides positive effects (e.g. increased productivity in business tasks (Cecchinato et al., 2015; Cousins & Robey, 2015; Duxbury et al., 2014; Fleck et al., 2015), tensions between work and family domains (Kreiner et al., 2009) can have a negative impact on an individual, resulting in stress or work and private domain overload (Kreiner et al., 2009). Individuals may lose control over their boundaries between work and private life domains (Jackson et al., 2006) resulting in a change from "work anytime and anywhere" to "work all the time and everywhere" (Cousins & Robey, 2015; Davis, 2002).

In the last decades, researchers have used boundary and border theory to analyze how individuals manage boundaries between work and family domains. Different boundary management tactics, styles and strategies have been developed (Allen et al., 2014). For example, Kreiner et al. (2009) describe different tactics priests use to leverage their technology in order to organize their boundaries within behavioral tactics. Findings of Duxbury et al. (2014) of the adoption and use of Blackberry smartphones indicate that successful boundary management depends on the development of a strategy in order to manage the device prior to adoption. However, research on technology related boundary tactics is sparse. Against this background, the objective of this study is to facilitate greater understanding of individual tactics to manage the boundaries between work and private life domains using information technology.

To answer this objective, the paper is structured as follows. First, we will define and describe the core themes of our study, namely boundary and border theory, and will explain how they have been used in general and in IS literature specifically (Section 7.2.2). After explaining our methodological approach (Section 7.2.3), we will present our findings in Section 7.2.4. In section 7.2.5, we will conceptualize and integrate our findings and discuss them in terms of potential generalization beyond our area of interest (Section 7.2.6). The paper concludes with an outlook, formulating the limitations as well as implications for future research and practice (Section 7.2.7).

7.2.2 Related Work

7.2.2.1 Boundary theory

Boundary theory (Ashforth et al., 2000; Clark, 2000; Nippert-Eng, 1996; Reyt & Wiesenfeld, 2015; Rothbard et al., 2005) refers to the way in which people try to create, maintain, change, simplify or order their environment. Specifically, boundary theory focuses on boundaries between roles. Katz and Kahn (1978) outline roles as expectation, placed on an individual in a social system. Therefore, in the context of our study we use the term boundary to describe a limitation of space and edge of a role, varying on a continuum from thin to thick (Allen et al., 2014; Kreiner et al., 2009). Thin boundaries are associated with being weak and open to influence, whereas thick boundaries are supposed to be strong and not influenceable (Ashforth et al., 2000; Hartmann, 1991).

Boundary theory has been used in different contexts e.g. psychology, organization theory and political science (Kreiner et al., 2009). Based on a cognitive theory of social classification with the focus on how people prioritize work and home (Allen et al., 2014) boundary theory evolved from sociological work of Nippert-Eng (1996). When applied to the work and family literature, boundary theory describes key challenges individuals face, managing work roles (e.g. as an employee) and family roles (e.g. as a parent) and the transition between those two roles, as they are defined as distinct from one another (Ashforth et al., 2000; Hall & Richter, 1988; Kossek & Lautsch, 2008; Nippert-Eng, 1996). The transition between those roles, as described above, can be of a psychological or physical way and can differ, regarding an individual's preference in terms of their flexibility and permeability (Ashforth et al., 2000). Due to the variance of transitions a continuum of border demarcation arises, showing on the one-hand integrators, (individuals, drawing a thin line between work and family roles) and on the other-hand separators (individuals, drawing a thick line between work and family roles) (Nippert-Eng, 1996). Ashforth et al. (2000) further distinguish between macro (infrequent, involving permanent change) and micro transitions (frequent, involving routine activities).

7.2.2.2 Boundary management – preferences, tactics and styles

Research of boundary theory states that there is a difference between boundary preferences, tactics and styles. Kreiner (2006) describes boundary preferences as an individual's preferences of either implementing or segmenting aspects of work and private life domains. An important aspect is that an individual's preference describes the wish of an ideal boundary management. Therefore, individuals use tactics to create their preferred style of segmentation or integration (Kreiner et al., 2009). Whereas the boundary preferences refer to the integration or segmentation preference, the boundary styles refer to the actual enactment of integration or segmentation (Kossek et al., 2012).

Kossek and Lautsch (2008) identified three different boundary management styles: integrators (blending work and family domains), separators (dividing work and family domains) and volleyers (switching between those two strategies). In order to define boundary management in more detail, different frameworks developed over time (Allen et al., 2014). Allen et al. (2014) identified two lines of research that arose based on Kossek and Lautsch (2008). One line identifies specific boundary management tactics (Kreiner et al., 2009) (Sturges, 2012) whereas the other line analyses boundary management styles (Ammons, 2013; Kossek et al., 2012).

Kossek et al. (2012) defined six different clusters that can be used to classify individuals that describe how an individual manage its personal preferences of boundary styles. These six clusters ("work-warriors", "overwhelmed reactors", "family guardians", "fusion lovers", "dividers" and "nonwork-electrics") differ regarding their control of demarcation, focus on work or family domains and break-behavior of boundaries (e.g. "fusion lover" and "nonwork-electrics" have a high control in contrast to "work warriors" and "overwhelmed reactors", whereas "fusion lovers" and "overwhelmed reactors" both focus on both work and family and "work warriors" and "nonwork-electrics" describe the ends of boundary continuums) – focusing on either work for "work warriors" or maintaining a small identification with their family for "nonwork-electrics". Break-behavior of "work warriors" is defined by a high permeation from work to private, whereas "overwhelmed reactors" are described by a break-behavior in both directions – work and family. "Fusion lovers" and "nonwork-electrics" tend to integrate both break-behavior patterns allowing work permeation during family time and the other way around (Kossek et al., 2012).

Since individuals are able to actively change their boundary style, Kreiner et al. (2009) describe tactics individuals use in order to design their preferred living of work-home integration and segmentation in daily life. These tactics can be of behavioral (e.g. involving other people), temporal (e.g. controlling work time), physical (e.g. managing separate artifacts for work and family domains) or communicative style (e.g. confronting boundary violaters either during or after a violation (Kreiner et al., 2009).

We carefully note that some work has been done in extant literature describing boundary management tactics using information technology. For instance Kreiner et al. (2009) describe a micro-category called "leveraging technology" which is a sub-category of behavioral tactics. This micro-tactic is linking directly to the use of information technology to manage boundary strategies. In his comprehensive study with Priests, they identify the use of voice-mail, caller ID, e-mail and the Palm Pilot Calendar as technologies that help them to facilitate their boundary management. Similarly, Duxbury et al. (2014) discovered individuals as not being able to segment between the two domains due to a lack of self-discipline and self-control when using smartphones (e.g. Blackberry). Koeffer et al. (2015) found six technology-related aspects (dual use of company IT for private task, dual use of private IT for work tasks, remote access to work data, distinct devices for private and work purposes, separate private and business accounts and quality of company provided IT), explaining the intensified professional use of IT. They concentrate on IT which was originally developed for the consumer market to manage boundaries between work and private life domains. Cecchinato et al. (2015) observe the use of e-mail accounts across devices to manage boundaries in more detail, finding micro-boundary strategies in e-mail management.

Although there has been significant research in the field of boundary management so far, only limited research addresses technological aspects on boundary management. Against the background of technological advancement including the emergence of IT Consumerization previous research show that technology influence boundary management (Koeffer et al., 2015). Consequently, more research is needed to shed light on technology related boundary management.

Therefore, we want to bridge this gap by further differentiating information technology micro-tactics. In order to identify these tactics, we conduct an explorative study with the objective to uncover IS tactics used by individuals to manage their boundary styles. Taking a qualitative approach, we build on the foundation of Kreiner et al. (2009), Kossek and Lautsch (2008), Koeffer et al. (2015) and Cecchinato et al. (2015) and extend current research by including technology related aspects. In order to address our aim, our research is guided by the following research question:

Research Question 1. How do individuals use IT in order to manage their boundaries between work and private life?

7.2.3 Research Method

Method selection. Although various studies from psychology and organizational science already explored and analyzed individual tactics and strategies to maintain boundaries, information systems research did not exploit the full potential of boundary theory so far. Therefore, this research pursues an explorative approach, to gain insights on how individuals use information systems to implement boundary management tactics. Based on the explorative nature of this study, we made use of tools from grounded theory methodology (Glaser & Strauss, 1967; Urquhart et al., 2010) which is explained next.

Data collection. We conducted a total of 15 interviews (10 males, 5 females). The participants were selected out of different organizations including industrial sector, financial sector, IT-business and public sector. An overview of the interviewees is presented in Table 7.8.

Position	No of Interviewees	Average work experience in years	Number of the interviews
Employee	9	6	1, 2, 3, 4, 5, 7, 8, 11, 13
Manager	6	14	6, 9, 10, 12, 14, 15

Table 7.8: Overview of Interviewees (Paper 10).

We conducted a two-step approach to conceptualize individual tactics. First, we conducted four semi-structured interviews. We included open questions like *"Do you separate private and business technology?"* or *"What are technological approaches to meet your boundary preferences?"* In this first round, we interviewed doctoral students from the business faculty (employees), because they are provided with mobile technologies and they have a great degree of freedom on how, when and where they work since they are generally managed by objectives.

Based on this first step, we further adapted our questions. We continued by interviewing another eleven individuals from industry. To get insights from different hierarchies, we included both employees and manager. Furthermore, we particularly included practitioners with working experience (9.2 years of working experience in average) to capture individual strategies that have been already implemented.

Data analysis. Following the grounded theory approach, we analyzed the data beginning with open coding (Corbin & Strauss, 1990; Glaser & Strauss, 1967). Three of the researchers implemented the procedure of open coding independently. They read the transcribed interviews and proposed codes that represent the content. Afterwards, similar codes were collected out of the interviews and grouped as a common denominator what is known as axial coding. For instance, for the subsequent citation *"I own an IPhone and it is equipped with the tool to only permit phone calls from people which I chose, at the times which I selected."* (Interview 1), three independent codes ("filtering", "manage communication", and "automatic filtering") were found. Finally, "filtering" was used as an axial code. Disagreements were discussed with the remaining researchers and settled by a mutual agreement.

We finished our process when all researchers agreed that there is only little chance that new essential concepts would emerge. Since our data highlights key aspects of the integration or separation between work and life, we finish our analysis by relating our results with existing literature (theoretical coding, Section 7.2.5).

7.2.4 Findings

Physical detachment. Kreiner et al. (2009) analyzed physical tactics describing dismantling local boundaries between work and private life domains. However, Kreiner et al. (2009) did not link physical tactics to IT. When looking at the interviews, we noticed that employees, having two devices, for example a private device and a corporate device, tend to separate between those two devices. Most commonly, they separate based on the ownership. Therefore, the corporate owned one is exclusively used for work and the private device in exclusively used for private purposes. The following excerpt illustrates this behavior:

> *"Ultimately, that's why I own two smartphones, one for work and one for my private matters. The same for computers. Generally, I respect the separation to use the company device only for work related issues and my private phone or laptop for everything else. [...] Well, that means, I keep the usage of my private device for company matters to the minimum. I would glance at emails via a SharePoint, but I would never download an Outlook Client to have fully access to my company emails." (Interview 12)*

For example, when looking at the private life domain, ways to foster separation using mobile devices could consist of leaving the corporate device at work, switching it off or to turn it to a silent mode. The following quote shows an individual separating using two ways. First, the silent mode is used in order to prevent interruption. Second, he puts the corporate smartphone aside in order to prevent a confrontation with checking it for notifications:

> *"After my working time, when I am at home or in the gym, I put my phone away – in silent mode- then I don´t realize that a message or a call came in and I won´t answer it." (Interview 15)*

Automatic notification. As technology enables the automatization of processes, it also opens the door for the individual boundary tactic, especially, in terms of communication applications there are prevailing ready-to-use configurations to define automatic notifications for instance in terms of absence times. A common use of automatic notifications can be found in E-Mail applications. The following excerpt describe how one employee use automatic E-Mail notifications.

> *"I assigned my email account to automatically answer received emails with the message "Thank you very much for your email, however right now I am unable to answer it, I will be back on XY-day." Obviously, after this email is sent and I return, I will check back to answer it appropriately. Then, of course, it will be my problem." (Interview 13)*

Although this excerpt illustrates how automatic notifications can be used, it also emphasize the importance of individual behavior. Conclusively, if an individual uses that tactic to separate, at this point, technology does not enforce a strict separation. *Pull information.* There are different ways of getting access to phone calls, e-mails and further information and notifications. Pulling information describes an individual's behavior to inquire their current notifications. One way is described as choosing where and when to get access to information and notifications. One employee describes his preference to pull e-mails from web account browser in order to be able to decide when and where to check e-mails:

> *"I determine the time. [...] That´s why I usually use the browser to access my emails. Using the online account, I decide when to check work emails." (Interview 13)*

Another employee states his preference on pulling information as viewing notifications on his smartphone, when turned on the silent mode, anytime and anywhere he prefers to:

> *"Most of the time, my private phone is in silent mode. Now and then, I would check if someone texted me and I would answer, although I am at work. It also depends on the moment, if I am very busy or if I have a little downtime to check my messages." (Interview 9)*

Pulling information is described by another employee as a routine defining when and where to check e-mails regarding, working together across different time zones: As different time zones implicate the possibility to get e.g. e-mails anytime, anywhere from everywhere, the employee talks about a routine behavior in order to cope with this permanent flow of information. He talks about a routine describing to pull information when you want to but to answer only if you need to:

> *"As I said, the time in China is 4 am when it is 10 pm here. On the other hand, it is 10 pm here in Germany when it is afternoon in the U.S. Since my company has offices everywhere, I could receive an email in the middle of the night. The message will be read, but by now, the routine is there." (Interview 10)*

Push information. Another way on getting information is not to decide when and where to access these information but rather just let these information go through anywhere and to anytime. In temporal intervals, e.g. e-mails being automatically queried, an individual gets to know new notifications using vibration or sounds to signalize these. An employee illustrates below how his e-mails are pushed anywhere at anytime:

> *"I receive every message. I don´t block out any notification. The internet on my phone is not shut down and I don´t disable private accounts, which I administer with my MacBook. That means, I am available all the time. However, whether I react to the notifications depends on the problem at hand." (Interview 7)*

Another employee states how she decided to get e-mails pushed at an interval of 30 minutes in order to be up to date with her notifications:

> *"Every half an hour I receive a notification. I assume half an hour is enough time, it doesn´t have to be adjusted to a minute-by-minute routine." (Interview 3)*

Different employees confirm that setting an automatic interval in order to get notifications about received e-mails is helpful to be all the time informed about work and private life domains happenings. It is also described as easier due to not to have to log in every time in order to be able to check for example their e-mails. An interviewee states below:

> *"I think that the email account is updated every 30 minutes. [...] I would have adjusted the settings similarly, to avoid logging in every time. However, this setup allows the emails to refresh automatically and I would have a look at the new emails." (Interview 4)*

Dynamic filtering. Employees who want to be available only for important issues when they are at work or at home have the opportunity to filter their incoming messages dynamically. When applying dynamic filtering, only messages or phone calls from specific individuals are received in a set time frame. For example, one employee explained that he told his smartphone to only let through phone calls from his family when he is at work.

> *"I own an IPhone and it is equipped with the tool to only permit phone calls from people which I chose, at the times which I selected. For example, from 10 am until 8 pm, only my family can reach me and they only call when it is important. All other callers are blocked. Like that, I created my own free time." (Interview 1)*

When using this tactic, employees mainly separate work and private life. They only want to integrate work and private life when an intrusion from the other domain is important enough for themselves. *Boundary App.* Technology can enable employees to manage their work life balance in helping them to focus on their currently active role. When employees are engaged in their work, technology prevents interruptions from family and private life. Similarly, when employees want to have private time, technology inhibits work related interruptions. Therefore, employees can integrate and separate to a certain degree to their own preferences. One employee illustrated this with a setting in his smartphone that enabled him to switch either to work or to private life:

> *"The new Blackberrys have a feature where you are able to separate work and your private information. That means, on one device you can switch between a work mode and a private mode. The private mode is used for private emails, WhatsApp, Facebook, etc. whereas work related emails can be checked using the work mode of the phone." (Interview 14)*

However, this technology might have both positive and negative effects. The advantage of a boundary app is that one can use the same device for multiple purposes without being interrupted from another life domain. Therefore, they can integrate their work and life at whatever time they like to but still keep this time free from interruptions because they separate. As a downside, at least in the context of our interviewee, there is the risk of invading users' privacy:

> *"The advantage is that I only have one device. However, the downside is that I give my employer information about my private life." (Interview 14)*

7.2.5 Conceptualization of Individual Tactics

The maturity of technology use is an important aspect with regard to our research question, because it has a major influence on how individuals implement boundary tactics. Maturity in general has been addressed in various IS studies for instance as an overall technological maturity (Karimi et al., 1996) or on an individual level based on self-efficacy (Venkatesh et al., 2003). Since we focus on individual tactics, self-efficacy and individual maturity in terms of technology use is most relevant. Automatization of business processes can be understood as a high level of maturity, whereby manual processes can be considered as low maturity (Dumas et al., 2013). Based on this distinction we propose four different domains of individual boundary tactics which are summarized in the following Table 7.9.

Boundary Preference	technological maturity	Implementation Tactic
Integration	High (automatic process)	Integration is integrated by automatic mechanisms (e.g. dynamic filtering)
	Low (manual process)	Integration is conducted loosely through manual mechanism (e.g. manual procurement of information)
Separation	High (automatic process)	Separation is implemented by automatic mechanisms (e.g. automatic response notifications)
	Low (manual process)	Separation is conduced manually (e.g. physical detachment)

Table 7.9: Four Domains of Individual Boundary Tactics (Paper 10).

Our findings suggest that there are various approaches to comply with the individual tactic. Since automatization of IT is often on a continuum (ranging from manual to full-automation), a strict separation is of these tactics is rarely possible. For instance, the configuration of a communication filter (e.g. disable phone-calls after 8 pm) has both manual and automatic parts. In that case, we would argue that the core mechanism, namely the filtering, is mainly automatic. Conclusively, we propose a matrix including a continuum from integration to separation (Ashforth et al., 2000) and a continuum describing the technological implementation from manual to automatic. Building on this framework, the domain-affiliation of the different tactics are summarized in Table 7.10.

Individual tactic	**Primary objective**	Examples for technological implementation
Physical detachment	separation	Leaving technology at work when at home; turning work-related technology off when at home or turning technology silent or on vibration.
Automatic response		Using an answering machine; sending e-mail-notifications for e-mails that arrive after hours or on vacation.
Pull Information	mediation between integration and separation	Actively looking up new messages and phone calls without being informed just in time.
Boundary App		Possibility to change actively within the same technology between home and private life domains.
Push Information	integration	Being informed just in time about incoming messages and phone calls.
Dynamic Filtering		Setting up filters that let notifications of specific individuals come through.

Table 7.10: Overview of Individual Tactics (Paper 10).

In summary, we identified six major IT tactics that allow individuals to maintain their boundary preferences. As they are located on a continuum (Ashforth et al., 2000), we recapitulate them in the following figure.

7.2.6 Discussion

Summary. Information technology fundamentally influences all aspects of our life. It is therefore not surprising that IT enables a multitude of possibilities to implement and maintain individual tactics to meet one's preferences. In order to answer our research

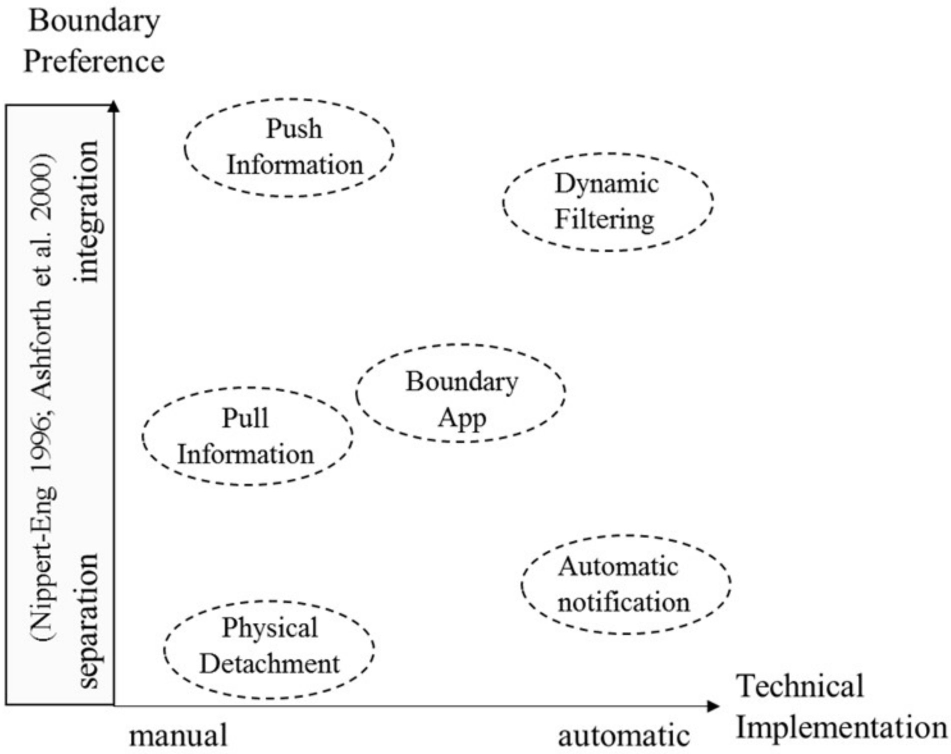


Figure 7.3: IT-related Boundary Tactics (Paper 10).

questions, we identified six different individual tactics (physical detachment, automatic notification, pull information, boundary app, push information, and dynamic filtering) and systematically categorized them with regard to boundary preferences and technical implementation (see Table 7.3).

Implications for theory. As our findings propose a more granular distinction of technology-related tactics, they enrich the findings of previous studies. By exploring individual boundary tactics, our research primarily contributes to boundary theory (Ashforth et al., 2000; Nippert-Eng, 1996). In particular, our findings enrich the boundary tactics from Kreiner et al. (2009) by differentiating technology-related tactics. As such we added another continuum dimension besides integration and separation, namely technological implementation, to include technology-related aspects based on their automatization level.

We also contribute to the study of Duxbury et al. (2014) who describe the complex relationship between mobile technologies and individual boundaries. Their results show that developing a strategy to manage the use of mobile devices across work and private life domains is essential for reducing conflicts between work and private life domains. Our findings can be further used to analyse the relationship between mobile technologies and boundary preferences against the background of the identified technological tactics (see Table 7.3).

Koeffer et al. (2015) suggest that there are six aspects related to the consumerization of IT that influence work-life balance. They propose that the allowance or the permission of these aspects leads to work-life balance and conflict. With our findings, we further develop this idea by proposing a set of alternatives that can be used to improve individuals balance (for instance by offering a “boundary app”). Finally, we also contribute to Cecchinato et al. (2015) who put emphasize on micro-boundary strategies related to e-mail accounts. By extending our research beyond e-mail communication, we further identified technology related aspects that are relevant for individual boundary management. Specifically, the use of a mobile “app” that is used for a broad variety of scenarios (e-mail, phone, text message etc.) allows valuable insight into individual strategies, that can be used to further develop the device management as proposed by Cecchinato et al. (2015).

Implications for practice. Based on our findings, we can derive implications for practice regarding the autonomy and the knowledge of the employee as well as the possibilities of the organization to influence an employee’s boundary management. First, since individuals have different preferences in general and in terms of boundary management it is recommended that organizations try to offer enough freedom to implement them. Related to technology this can be done by offering chances to adapt and personalize technology.

Second, an individual’s knowledge on technology is a main aspect on implementing boundary preferences. Without sufficient capabilities to adapt technology, individuals are not able to meet their preferences. According to person-organization fit (Chatman, 1989; French et al., 1982; Kristof, 1996) organizations are encouraged to further train their employees on how to use (mobile) technology with a focus on individual adaptation.

Finally, organizations can easily influence an individual’s boundaries by setting defaults. For instance, when using a pull mechanism as default for e-mail communication, it is most likely that a great number of employees do not change to push (Thaler & Sunstein, 2009). Therefore, the organization can facilitate separation between private and work life.

7.2.7 Limitations and Outlook

Limitations. Besides common limitations of qualitative research, this study has limitations that are worth mentioning. First, we asked the interviewees about general tactics related to IT. However, in specific scenarios, for instance employees using wearables or augmented reality technologies which can be even less separated in terms of boundaries than mobile technologies, there might be more tactics which we did not cover so far.

Furthermore, using the level of technology automation is only one possible dimension with regard to technology. Others could be mobility, complexity or ubiquity. Therefore, our findings are limited to only one specific dimension. However, our findings are well suited to transfer to other dimensions as well.

Outlook. As our study explored general tactics with regard to boundary management, our findings propose a sound foundation for future research. Especially with regard to design science, experimental research could further explain various effects by matching individual preferences and the design of IT artifacts. Furthermore, affective technologies can be included in order to be able to identify individual's preferences.

7.2.8 Acknowledgements

This paper was written in the research project SUGAR (promotional reference 01FA15003) which is funded by the German Federal Ministry of Education and Research (BMBF).

7.3 Paper 11: Towards an Explanatory Design Theory for Context-dependent Learning in Immersive Virtual Reality

Title	**Towards an Explanatory Design Theory for Context-dependent Learning in Immersive Virtual Reality**
Authors	Katharina Jahn katharina.jahn@uni-siegen.de Henrik Kampling henrik.kampling@uni-siegen.de Hans-Christian Klein christian.klein@uni-siegen.de Yasin Kuru yasin.kuru@students.uni-siegen.de Bjoern Niehaves bjoern.niehaves@uni-siegen.de
Publication Type	Conference Proceedings
Publication Outlet	Pacific Asia Conference on Information Systems
Outlet Information	JOURQUAL3: C
Status	published
Full Citation	Jahn, K., Kampling, H., Klein, H.-C., Kuru, Y., & Niehaves, B. (2018). Towards an explanatory design theory for context-dependent learning in immersive virtual reality. *Proceedings of the 22th Pacific Asia Conference on Information Systems (PACIS 2018), Yokohama, Japan*

Table 7.11: Fact Sheet Paper 11

Towards an Explanatory Design Theory for Context-dependent Learning in Immersive Virtual Reality

Abstract

Immersive virtual reality (IVR) is increasingly used for learning. However, research on specific designs for IVRs which can be used to enhance individual learning performance is still at an early stage. In this research-in-progress paper, we build upon theories on context-dependent learning to develop an explanatory design theory. We hypothesize that if the user learns in a virtual environment that represents the recall environment (environmental congruence), recall is facilitated. Additionally, if the IVR is designed with a high degree of sensory immersion, the effect of environmental congruence on learning is further increased through enhanced cognitive absorption in the technology. In contrast, cognitive absorption in the task should have a reversed effect. To test the explanatory design theory, we plan to conduct a 2 (learning environment: Room A vs. Room B) x 2 (sensory immersion: low vs. high) x 2 (recall environment: Room A vs. Room B) between-subjects laboratory experiment.

Keywords— Immersive virtual reality, cognitive absorption, context-dependent learning, place-dependent learning, explanatory design theory, design science, laboratory experiment

7.3.1 Introduction

Forms of immersive virtual reality (IVR), a technology in which the user is completely absorbed into by the use of head-mounted displays, are increasingly used for learning in different contexts. There are IVR applications used for learning in schools, universities and in health care (Martín-Gutiérrez et al., 2017). Additionally, organizations such as VW started to use IVR for letting their employees learn new organizational processes (RoadToVR, 2018).

IVR has not only the advantage that learning can be designed in a way that is highly engaging, it also can be used to re-create places that are not easily available to the learner. For learning, the latter can be especially beneficial because of environmental context-dependent memory. According to research on environmental context-dependent memories (Isarida & Isarida, 2014), learning and recalling in the same place is more beneficial for individual learning performance than learning and recalling in different places. However, research has not yet investigated if and how environmental context-dependent memory effects can be recreated by using virtual learning environments for recall in a real physical place. With the research-in-progress paper at hand, we therefore would like to address the following research question.

Research Question 1. How can IVR be designed to enhance context-dependent learning when it is not possible to learn in the environment where recall takes place?

Explanatory design theories (Baskerville & Pries-Heje, 2010; Gregor, 2009; Kuechler & Vaishnavi, 2012; Niehaves & Ortbach, 2016) can provide a framework to answer this research question by not only stating how to design an artifact, but also explain why specific design options have specific effects (Gregor, 2009; Kuechler & Vaishnavi, 2012) through the use of structural equation modeling terminology (Niehaves & Ortbach, 2016). In this research-in-progress paper, we draw upon theories on environmental context-dependent memory and cognitive absorption (CA) to develop an explanatory design theory for context-dependent learning that answers our research question (see Figure 1). We propose that environmental congruence enhances individual learning performance and that this effect is further increased by a high degree of sensory immersion which increases CA in the technology through heightened presence. We plan to conduct a 2 (learning environment: Room A vs. Room B) x 2 (sensory immersion: low vs. high) x 2 (recall environment: Room A vs. Room B) between-subjects experiment to test the explanatory design theory.

7.3.2 Theoretical Background and Model Development

In this section, we explain theories on environmental context-dependent learning and draw upon theories on cognitive absorption to relate environmental context-dependent learning to virtual reality (see Table 7.12 for an overview of construct definitions). On the basis of these theories, we develop our research model regarding the relationship between sensory immersion, environmental congruence, cognitive absorption in task and technology, and individual learning performance (see Table 7.4 for a hypotheses overview in the proposed model).

7.3.2.1 Environmental Context-Dependent Learning

Theories about context-dependent memory (Isarida & Isarida, 2014; Smith & Vela, 2001) state that different contextual cues can affect recall of target information. Whereas target information is defined as the information that should be remembered, contextual cues represent information that is not the target information but was present (physically or mentally) during encoding. If the context is encoded with the target information,

Construct	Definition	Source
Cognitive Absorption in Task (CA_TASK)	Cognitive absorption is defined as an enjoyable state of deep (cognitive) involvement in the *performed task*	(Csikszentmihalyi, 1990)
Cognitive Absorption in Technology (CA_TECH)	Cognitive absorption is defined as an enjoyable state of deep (cognitive) involvement with the *technology used.*	(Agarwal & Karahanna, 2000)
Telepresence	Telepresence refers to perception of the user in contrast to the technology design. It is defined as the degree to which an individual perceives to be in a distant place.	(Schultze, 2010) (Schultze, 2014)
Sensory Information	Sensory immersion describes the design of the technology in contrast to the perception of the user. It is defined as the degree to which a technology can achieve convincing illusion of reality to the users' senses.	(Schultze, 2010) (Slater & Wilbur, 1997)

Table 7.12: Construct Definitions (Paper 11)

the context can be used as retrieval cue for remembering (Isarida & Isarida, 2014). Therefore, the learning performance of individuals (individual learning performance, ILP) can be enhanced through the use of context-dependent learning.

Theories about environmental context-dependent memory specify this effect for aspects related to the environment in which the target information was learned (Isarida & Isarida, 2014). Environmental context can consist of the larger environment, such as place (Smith and Vela 2001) or specific aspects of the environment, such as odor (Isarida et al., 2014), background color (Isarida & Isarin, 2007), or background music (Isarida et al., 2017). For example, when an employee learns how to use a machine, new to the organization in an office, different elements of the learning environment (e.g. lightning, desktop computer, background music) can be encoded with the target information during learning. Therefore, recall of the target information might be hindered, when the employee tries to remember it in a different learning environment, such as a production hall, which consists of different environmental elements.

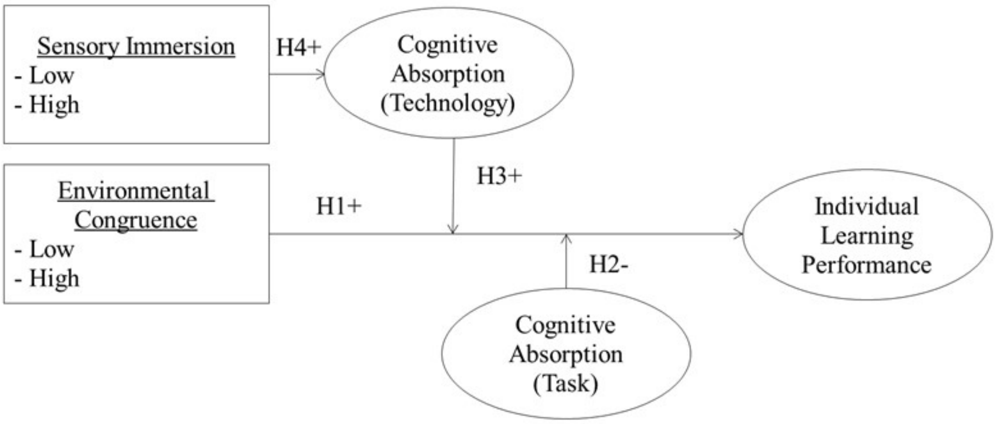


Figure 7.4: Explanatory Design Theory for Context-Dependent Learning (Paper 11).

With physical reinstatement of the environment in which the learning took place, recall is facilitated (Godden & Baddeley, 1975; Isarida & Isarida, 2014; Smith et al., 1978). In the case of the employee, if they return to the office, the physical reinstatement of the environmental context could enhance recall, because the elements of the environmental context were encoded together with the target information. However, physical reinstatement is often difficult in practice. Returning to the office for recall every time the machine gives an error message which the employee learned at the office would be time-consuming. Additionally, returning to the office might not be a helpful solution if the employee has to remember the information in the production hall while using the machine and not in the office. Therefore, an alternative to physical reinstatement would be mental reinstatement – the mental visualization of the environmental learning context. However, mental reinstatement can be too difficult in some circumstances (Canas & Nelson, 1986) or needs to be requested explicitly in order to be used for some individuals (e.g. older adults, (Fernández & Alonso, 2001).

With the use of IVR, it is possible to simulate the context almost completely without being physically located in the initial learning environment. Thus, with the use of IVR in the learning situation, employees could benefit from the advantages of physical reinstatement without the costs of mental reinstatement. However, research on using IVR in the learning situation and recall in the actual physical environment has not yet been done. Based on the described place-dependent memory effects in real learning and recall environments, we assume that learning and recall in a congruent environment is more beneficial to the user than learning and recall in an incongruent environment. Regarding the design, we therefore hypothesize that learning in a room that is congruent to the recall room is more beneficial for recall than learning in a room that is incongruent to the recall room.

Hypothesis 1. Environmental congruence leads to a higher ILP than environmental incongruence.

7.3.2.2 Cognitive Absorption

In the field of learning, cognitive absorption (CA), also called flow, refers to a state in which an individual is completely involved with a task (CA_TASK, (Csikszentmihalyi, 1990)). CA_TASK is usually identified to be a desirable state for learning (see Table 7.12 for an overview of construct definitions). However, some research on CA_TASK suggested that high levels can be detrimental because context effects are blocked out (Magni et al., 2013). Likewise, research on environmental context-dependent learning has indicated that a high involvement with the task, and therefore high CA_TASK, decreases ILP because the environment is blocked out (Smith & Vela, 2001). We therefore hypothesize that CA_TASK moderates the effect of environmental congruence on ILP.

Hypothesis 2. CA_TASK moderates the relationship between environmental congruence and ILP. For individuals with a low level of CA_TASK, the relationship between environmental congruence and ILP will be higher than for individuals with a high level of CA_TASK.

In the field of Information Systems, CA has been conceptualized as the state of being completely immersed in a technology (CA_TECH, (Agarwal & Karahanna, 2000; Burton-Jones & Straub, 2006), letting the role of the context strongly depend on the technology referred to. For example, in the study of Agarwal and Karahanna (2000), the technology in which an individual was cognitively absorbed in was the web, whereas Burton-Jones and Straub (2006) referred to MS Excel. If these constructs initially developed in the context of technology acceptance are adapted in the field of learning, it is important to note the different implications a high CA_TECH might have in both cases. If an individual has to learn something in the web, a high CA in the web does not necessarily imply a high CA_TASK because the web can be used in a range of task-unrelated ways very easily. In contrast, MS Excel still can be used in task-unrelated ways while being highly cognitive absorbed (e.g. drawing pictures instead of calculation), but the affordance for these alternatives is probably much lower than in the case of the web.

In the context of IVR, CA_TECH leaves the user even more room for task-unrelated activities. By sealing the participants from the actual world through a head-mounted display and earphones, an almost completely immersing new virtual world is created. Therefore, context that would traditionally be neither part of the task nor part of the technology (e.g., a cupboard displayed in the IVR) becomes a part of the technology. Thus, the meaning of CA_TECH changes dramatically in IVR by covering a much broader range of the environment.

Whereas the described unspecificity of CA_TECH is not that important for studies of technology acceptance, it needs to be addressed in the area of learning because of the confusion with CA_TASK. Studies that have used items that resembled CA_TECH instead of CA_TASK in the learning context showed that CA_TECH might enhance learning through a motivational route by affecting learner satisfaction (Leong, 2011) and continued use (Guo et al., 2016) as well as perceived learning (Reychav & Wu, 2015). However, these studies did not vary cognitive absorption experimentally and used technologies such as computers, smartphones or tablets instead of IVR. For the relationship between CA_TECH and learning in an IVR, a qualitative research gives initial support for a relationship between CA_TECH and learning (Kampling, 2018). Therefore, we want to address this research gap and investigate whether CA_TECH has an influence on actual (instead of perceived) learning outcomes for declarative knowledge.

In the field of context-dependent learning in an IVR, CA_TECH might influence the relationship between environmental congruence and learning. We assume that a higher CA_TECH before the learning task will lead to a stronger encoding of contextual information which can then strengthen the relationship between environmental congruence and learning.

Hypothesis 3. CA_TECH moderates the relationship between environmental congruence and ILP. For individuals with a high level of CA_TECH, the relationship between environmental congruence and ILP will be stronger than for individuals with a low level of CA_TECH.

7.3.2.3 Immersive Virtual Realities and Cognitive Absorption

IVR can enhance the sense of “being there” – usually called telepresence – (Schultze, 2010, 2014) by presenting a high degree of sensory immersion to the user. Whereas telepresence refers to the psychological perception of the user, sensory immersion refers to the objective criteria of the technology design. Sensory immersion is therefore defined as the degree to which a technology can achieve an inclusive, extensive, surrounding and vivid illusion of reality to the users' senses, matches the users' movements to the visualizations of the IVR, and presents a convincing plot to the senses of the user (Slater & Wilbur, 1997).

Different factors of sensory immersion influence telepresence positively (Cummings & Bailenson, 2016) which in turn is positively related to CA_TECH (Faiola et al., 2013). High sensory immersion should therefore lead to higher telepresence and CA_TECH than low immersion. We therefore hypothesize an interaction effect of sensory immersion and environmental congruence on ILP, which is mediated by CA_TECH for sensory immersion.

H4a: There is an interaction effect of sensory immersion and environmental congruence on ILP. High sensory immersion strengthens the effect of environmental congruence on ILP more than low sensory immersion.

H4b: The interaction effect of environmental congruence and sensory immersion is mediated by CA_ TECH for sensory immersion.

7.3.3 Method

7.3.3.1 Participants and Design

We plan to recruit 200 students of Information Systems and Business for the experiment who receive a compensation of 5€ for their participation. We use a 2 (learning environment: Room A vs. Room B) x 2 (immersion: low vs. high) x 2 (recall environment: Room A vs. Room B) between subjects design. We rely on Information Systems and Business students as participants because with them, we can let them take the IVR experience at a physical location where they learn frequently (the IS department). For the recall setting, we can then use one place in which they have never been (the research center) and one place in which they are only occasionally (the main university). By doing this, we will be able to create a similarity to the situation in which the user wants to learn in an environment which they can visit only with difficulty and therefore does not necessarily visit it often.

7.3.3.2 Materials

Hardware and Software. The entire virtual environment is designed with the game engine Unity and the use of a 360° camera as well as the use of 3D laser scanning for the two rooms in which the learning takes place. The use of the 3D laser scanning makes sure that the participants can walk freely in the room and sit on a chair and at a table that are modeled in accordance with the real ones in both contexts. The chair is tracked with a HTC Vive Tracker to allow participants to sit down without falling. All participants wear a head-mounted display (HTC Vive) for viewing the IVR in the learning phase. In front of the HTC Vive, the Leap Motion technology (similar to Schwind et al. (2017)) is mounted for all participants, even though it displays the tracked hands into the virtual scene in real-time only for participants in the high sensory immersion condition. Additionally, we let the participants in all conditions wear three HTC Vive trackers (two on each foot and one on the hip) for full-body tracking with Ikinema Orion which are also only functional for participants in the high immersion condition . For the audio aspects within the experiments, a noise cancelling headphone is used. At the beginning, each participant is fitted with the headphone and active noise canceling.

Learning Task. Comparable to similar studies used for context-dependent learning (Godden & Baddeley, 1975; Smith et al., 1978), we use a word list consisting of 40 common, four-letter words that the participants have to remember. The words are presented via headphones and the space between words consists of an interval of 3 seconds.

Sensory immersion. In the low sensory immersion condition, participants wear a head-mounted display, controllers to interact with the virtual environment, and a headphone through which no sound is played. Instead of having a body, participants only see two controllers with which they interact in the virtual world. In the high sensory immersion condition, participants wear a head-mounted display and, using Leap Motion, they can interact with the virtual environment using their hands which are displayed through Leap Motion in the VR. Additionally, they can see a body when they look down which is tracked through the HTC Vive trackers. Background music is played through the headphones, different for each context.

Contexts. Both contexts are presented virtually in the laboratory of the Information Systems department of the local university for the learning phase and are later visited physically for the recall phase. The contexts differ in how the two rooms look and where they are located. Room A is located at the research center of the local university which is about 15 minutes by bus from the Information Systems department. Room

A is designed similar to an office and participants sit at a table on which a desktop computer, a telephone and various office tools stand. They look at a flip chart and a cupboard filled with books. A specific background music is played in the room. Room B is located at a building in the main university which takes about 20 minutes by bus from the Information Systems department and about 10 minutes by bus from the research center. The room is associated with a specific background music consisting of different classical music pieces. Room B is similar to an office and participants sit at a table on which a desktop computer, a telephone and various office tools stand. They look at a flip chart and a cupboard filled with books. A different background music consisting of other classical music pieces is played for Room B with the same tonality and tempo (similar to Isarida et al. (2017), who used a design in which background music did not affect mood).

7.3.3.3 Procedure

The experiment is divided in two sessions, the first one for the learning phase and the second one for recall. The first phase, were the exposition to the IVR-setting takes place, is located at the Information Systems department of the local university. Participants are tested individually. When they enter the laboratory, the experimenter tells them that the experiment is about experiences in VR and explains them how to put on the head mounted-display and the trackers. After participants have put on the head-mounted display, they see the outside of the building in which Room A or Room B is located in a 360° video. Then they see a virtual walk through the door of the building and enter it. In the building, they walk to the door of either Room A or Room B (depending on the building). Participants are then instructed by headphones to open the door either with the controllers in the low immersion condition or with their hands in the high immersion condition. They can then walk freely towards the chair to sit at the table. After they answer the presence and CA_TECH questionnaires, they are told that they hear a word list and that they should try to remember the words. Participants then hear the 40-word word list. Afterwards, similar to the procedure of Smith et al. (1978) the word list is presented again and participants have ten seconds between each word to rate the affective value of each word on a continuum from "good" to "bad" using either the controllers (low body tracking) or their hands (high body tracking). We use this approach to induce a sense of closure for the session and prevent participants from rehearsing the list between sessions. Participants then answer the questionnaire for CA_TASK. Participants are told that they should come to either Room A or Room B on the next day at a specific time in order to answer a final questionnaire and to receive their compensation fee.

The second session takes place about 24 hours later and is located either at the research center of the university (Room A) or a building in the main university (Room B). When they arrive at the room, the experimenter explains to the participant that they should write down as many words as they can remember in a surprise free recall test. The experimenter then leaves the room for 10 minutes. Afterwards, subjects are asked whether they have rehearsed any words between sessions, fill out the questionnaire of perceived room similarity between learning and recall room, answer questions for perceived learning, receive their compensation fee and are thanked and debriefed.

7.3.3.4 Measures

Individual learning performance. Individual learning performance is measured by the number of items recalled and by a perceived learning questionnaire adopted from Magni et al. (2013).

Cognitive absorption. We adapt the 5-item-measure of Burton-Jones and Straub (2006) for CA_TECH and CA_TASK. We frame the CA_TECH items towards the technology, similar to Burton-Jones and Straub, and the CA_TASK items towards the task, similar to Magni et al. (2013).

Manipulation checks. We use the Igroup Presence Questionnaire (Schubert et al., 2001) as manipulation check for Immersion and questions for perceived room similarity as manipulation checks for environmental congruence.

7.3.4 Discussion

With the experiment, we plan to show that environmental congruence can be designed in IVR and that is enhanced by sensory immersion through CA_TECH. By letting participants learn a word list, an approach that is often used in basic research on memory, we want to show that the proposed explanatory design theory can be used for a range of different tasks. Whereas the relevance for practice would have been more obvious with a task that focused on application in an organization, recall of a word list represents a basic function of memory in general. Therefore, the underlying mechanisms of the explanatory design theory should apply for most tasks in which recall of declarative knowledge is relevant.

On the basis of our results, we expect that future research can extend our explanatory design theory to different types of knowledge (e.g. implicit knowledge), compare it with the effects in reality, and use it as basis for identifying additional design options to enhance ILP. Specifically, as we operationalized sensory immersion by varying different design features (tracking level, the use of an avatar and background music) to enhance the feeling of being in a specific place, future research can identify whether there are differences in the role they play for environmental context-dependent memory or if there are design options that are more relevant than the ones we selected.

7.3.5 Acknowledgements

This paper is part of the research projects ELISE (promotional reference 16SV7512, German Federal Ministry of Education and Research (BMBF)) and ANTARES (promotional reference 005-1706-0006, Ministry for Culture and Science of the State North Rhine-Westphalia).

7.4 Paper 12: Designing for Knowledge-Based Familiarity, Trust, and Acceptance

Title	**Designing for Knowledge-Based Familiarity, Trust, and Acceptance**
Authors	Katharina Jahn katharina.jahn@uni-siegen.de Oliver Heger oliver.heger@uni-siegen.de Henrik Kampling henrik.kampling@uni-siegen.de Krzysztof Stanik krzysztof.stanik@uni-siegen.de Bjoern Niehaves bjoern.niehaves@uni-siegen.de
Publication Type	Conference Proceedings
Publication Outlet	European Conference of Information Systems
Outlet Information	JOURQUAL3: B
Status	published
Full Citation	Jahn, K., Heger, O., Kampling, H., Stanik, K., & Niehaves, B. (2017). Designing for knowledge-based familiarity, trust, and acceptance: The case of affective technology. *Proceedings of the 25th European Conference on Information Systems (ECIS 2017), Guimarães, Portugal*

Table 7.13: Fact Sheet Paper 12

Designing for Knowledge-Based Familiarity, Trust, and Acceptance

Abstract

With the ability to recognize human emotions, so-called affective technology has the potential to provide highly adaptive service to its user in many different areas such as learning, health care, or manufacturing. However, there are specific barriers for the acceptance of affective technology because most people are unfamiliar with the affective components of such technologies and, hence, do not trust them. Assuming that increasing the knowledge-based familiarity with an affective technology is essential for accepting it, so far, only little is known about appropriate design concepts to increase the familiarity and, as a consequence, the acceptance of affective technology. To close this gap, we follow a Design Science approach laying out an explanatory design theory for knowledge-based familiarity and acceptance of affective technology. We argue that familiarity with a technology is built by gaining knowledge about the emotional state the system has recognized and the subsequent behavior of the system and such knowledge will be gained by providing suitable feedback. We develop different designs for feedback systems of an affective technology and propose corresponding design hypotheses. This research-in-progress concludes with the planned experimental approach varying feedback content and feedback explanation.

Keywords— Affective Technology, Trust-based Acceptance, Knowledge-Based Familiarity, Feedback, Design Theory, Experiment

7.4.1 Introduction

With the ability to recognize human emotions, affective technology has the potential to provide situationally and individually highly appropriate service to its user. Since Picard's groundbreaking book (Picard, 1997), the field of "affective computing" has established itself dealing with the research and development of different affective technology applications, such as in education, security, health care, entertainment, marketing, and many more (Afzal & Robinson, 2015; D'Mello & Calvo, 2013). For instance, affect-aware learning technologies can detect boredom, confusion, frustration, or engagement of the learner based on conversational cues, body language, and facial features and respond adequately to improve the learning experience and to increase the learning effect (D'Mello & Graesser, 2013). In accordance with the definition of affective computing by Picard (2015), we define affective technology as technology which can sense and/or generate human emotions such as happiness, anger, or fear.

The acceptance of affective technology, that is the intention to use an affective technology application, is a key condition to make use of the service that such a technology can provide. Nevertheless, so far, research on the acceptance of affective technologies is rather fragmentary. In a qualitative study, Heger et al. (2016) have found that trust, knowledge-based familiarity, and emotional self-reflexivity are key conditions for the acceptance of affective technology. More research on the acceptance of affective technology in a broader sense deals with ethical issues (e.g. Cowie (2015). Assuming that trust as an antecedent of acceptance and knowledge-based familiarity as an antecedent of trust are essential for the intention to use affective technology (Heger et al., 2016), only little is known about how to make use of this to develop appropriate design concepts to increase the acceptance of affective technology. We therefore identify a missing design theory for the acceptance of affective technology as a research gap. Hence, the research in progress paper at hand has two research objectives:

> RO1: We develop and present a design theory for trust-based acceptance of affective technology.
>
> RO2: We outline the research method with which we will test the proposed design theory.

To achieve our research objectives, we develop an explanatory design theory (Baskerville & Pries-Heje, 2010; Gregor, 2009; Niehaves & Ortbach, 2016) comprising the constructs of behavioral intention, trust, knowledge-based familiarity, emotional self-reflexivity, and feedback. We argue that familiarity with a technology is built by gaining knowledge about the emotional state the system has recognized and the subsequent behavior of the system and such knowledge will be gained by providing suitable feedback. We develop different designs for feedback systems of an affective technology and propose corresponding design hypotheses. To test the developed hypotheses, we propose an experimental approach based on the application of an affective learning system.

7.4.2 Related Work

7.4.2.1 Acceptance of Affective Technology: Theory Background

In a qualitative study, Heger et al. (2016) identify trust, understanding the behavior of an affective technology (i.e. knowledge-based familiarity), and emotional self-reflexivity as most important for the acceptance of affective technology. Further research studies on affective technologies with regard to acceptance deal with ethical or social issues. Ethical concerns mentioned by Picard (1997) and Picard (2003), Reynolds and Picard (2004), and Cowie (2015) refer to the risk that emotion-related data are ultimately private and personal information that can potentially be provided to third parties. In contrast, data protection and considered privacy concerns could support acceptance and rise trust Picard (2003). Cowie (2015) summarizes the relationship between ethics and affective technology to the needs of "characteristic imperatives: to increase net positive affect, to avoid deception, to respect autonomy, to ensure that system's competence is understood and to provide morally acceptable portraits of people" (p. 334).

Trust in IT has been examined in several studies (e.g., Gefen (2000), Gefen et al. (2003), Komiak and Benbasat (2006), and McKnight et al. (2011). Rousseau et al. (1998) define trust as "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another" (p. 395). High social complexity arises when parties do not always behave rationally and predictably and, yet, people seek to understand them. This complexity can be reduced by trust (Luhmann, 1979; Rousseau et al., 1998). In addition, trust depends on the belief on an individual that the other party behaves dependably (Kumar, 1996; Kumar et al., 1995), ethically (Hosmer, 1995), and in a socially appropriate way (Zucker, 1986). In the context of Information Systems, McKnight et al. (2011) state that trust is not only relevant for person-to-firm relations and interpersonal relationships, but that "trust in the information technology itself plays a role in IT-related beliefs and behavior" (p. 1). Complementarily, Lewicki and Bunker (1995) argue that the other party can be an individual or an object. Moreover, Fukuyama (1996) states that trust is an essential and necessary precondition for the acceptance and adoption of unpredictable, uncontrollable, hazardous, and new technologies.

Trust arises by being familiar with the what, who, how, when, and what is happening (Lewicki & Bunker, 1995). Knowledge-based familiarity requires knowledge and information about the other and relies on the predictability of their behavior (Lewicki & Bunker, 1995). Knowledge-based familiarity reduces uncertainty by creating an understanding for what is happening in the present (Gefen et al., 2003; Luhmann, 1979). Gefen (2000) defines knowledge-based familiarity as a "specific activity-based cognizance based on previous experience or learning of how to use the particular interface" (p.727). In addition, Doney et al. (1998) state that trust arises from a prediction process based on knowledge and information as well as the anticipation of the other's party behavior.

7.4.2.2 Design of Feedback Systems of Affective Technologies

Feedback has fundamental influence in a broad variety of settings, such as in learning, working, or training environments. Feedback is an information delivery mechanism used to evaluate the extent to which prior behavior of an individual meets their internal goal standard (Martocchio & Webster, 1992). In the context of work, feedback is directing, reinforcing and shaping people's subsequent behaviors and performance within formal organizations (Moon & Sproull, 2008). In the context of training, feedback is meant to improve the employees' performance and to implement procedures (Ridder et al., 2015). The benefits of feedback, for instance, in learning environments has been examined in different experimental settings (Graesser et al., 2005), and in a variety of studies (Arnold et al., 2006; Butler et al., 2013; Kulhavy & Stock, 1989; Smits et al., 2008). Graesser et al. (2005) differentiate between positive, neutral, and negative feedback, which can be delivered from the system through prompting the user to fill missing information – through hints for how to go further, corrections if the input is wrong, and assertions as well as summarizations if information is missing or need to be shorten (Graesser et al., 2005). The main goal of feedback here is to deliver the possibility of adjusting actions towards the desired outcome. One crucial element for the design of contextual feedback is the information content (Butler et al., 2013).

Within the affective technology literature, there are many examples for designing feedback for emotional states recognized by the system. Video feedback is used in healthcare systems (Stratou et al., 2015), visual robots in the e-learning context (Wenhui et al., 2009), and visual emotional avatars in social learning environments for people with schizophrenia (Bekele et al., 2017), or in gaming settings (Sourina & Liu, 2013). Kummer et al. (2012) present an approach that reflects emotions of conversational partners by playing music. The studies of Landowska (2013) and Hupont et al. (2013) examine multi-methodical approaches. The first one uses text, audio, and video for tutoring systems while the latter uses smileys, emotional saccade (paths) maps, heat maps, and dashboards for reflecting affective states. In addition, Katmada et al. (2015) use graphs including a timeline in a serious gaming context. Carvalhaes et al. (2013) present a real robot called MollyPet which recognizes emotions of autistic children for therapy purposes. Moreover, Kerr and Bornfreund (2005) develop the virtual and emotional "BuddyBot" to generate consumer trust on a website.

In summary, there already is a substantial body of knowledge in the areas of trust and knowledge-based familiarity (in relation to technology acceptance), designing feedback in general, and designing feedback in the context of affective technology. Moreover, some research has been done in regards to the acceptance of affective technology. Nevertheless, knowledge about how to design appropriate feedback to increase familiarity with an affective technology application is missing. In addition, to our best of our knowledge, no investigation yet has studied the relation between the design of feedback and the construct of knowledge-based familiarity with a technology. We identify both as research gaps which this study wants to address.

7.4.3 Research Model and Hypothesis Development

In this section, we derive hypotheses to develop an explanatory design theory (Baskerville & Pries-Heje, 2010; Gregor, 2009; Kuechler & Vaishnavi, 2012; Niehaves & Ortbach, 2016). In contrast to design practice theories in which theory development is aimed at informing practice how to design artifacts, explanatory design theories are aimed at "analysing, describing and predicting what happens as artifacts exist and are used in their external environment" (exterior mode; Gregor, 2009, p. 7). For explanatory design theories, this aim can be achieved through setting up hypotheses that can be tested empirically (Gregor, 2009; Niehaves & Ortbach, 2016). Thus, on the basis of structural equation modeling language, an inner and outer model can be derived for design theory testing (Niehaves & Ortbach, 2016). Figure 7.5 represents the research model.

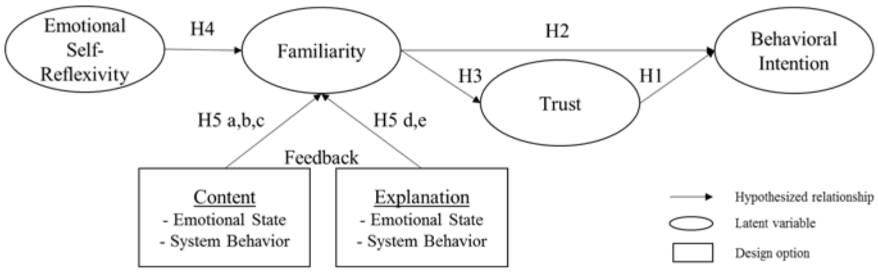


Figure 7.5: Research Model (Paper 12)

Acceptance of affective technology depends on trust in affective technology due to the fact that affective technologies are unknown to most people, are less controllable than common technologies, and operate with highly sensitive data (Heger et al., 2016). The studies from, for instance, Gefen et al. (2003), Gefen (2000), and McKnight et al. (2011) show that trust is especially important in situations with high uncertainty, which is the case when using affective technology. Moreover, Gefen (2000) argues that familiarity has a direct influence on behavioral intention for the reason that people who are overwhelmed by the complexity of an interface are more likely to give up using it.

H1: Trust in affective technology increases the behavioral intention to use affective technology.

H2: Knowledge-based familiarity with affective technology increases the behavioral intention to use affective technology.

Gefen et al. (2003) and Gefen (2000), for instance, show that knowledge-based familiarity is an antecedent of trust. Heger et al. (2016) propose that, in the context of affective technology, understanding an affective technology is significantly important for trust, since the fear of arbitrary behavior by the system will decrease if the behavior of the system is comprehensible, predictable, and can be anticipated.

H3: Knowledge-based familiarity with affective technology increases trust in affective technology.

According to Heger et al. (2016), "understanding affective technology depends on the capability to reflect and to be aware of one's own emotions", because "only by comparing the technology's behavior with their own perception, a user can develop an understanding of how and why an affective technology behaves in a certain way" (p. 8). They state that the concept of emotional self-reflexivity overlaps with studies and concepts from psychology, such as "emotional intelligence" (Dulewicz & Higgs, 2000), "self-awareness" (Steiner, 1997), and "emotional competence" (Bar-On & Parker, 2000; Boyatzis et al., 2000; Ciarrochi & Deane, 2001). The sub-constructs "comprehension", "clarity", and "awareness" of emotional competence developed by Goleman (1999) are said to highly overlap with emotional self-reflexivity (Heger et al., 2016). We add the category of "sensations" as an affective technology recognizes emotions through physical measurements.

H4: Higher levels of emotional self-reflexivity corresponds to higher levels of knowledge-based familiarity with an affective technology.

With the objective to increase a user's familiarity with an affective technology, the integration of a feedback system can be a solution. As the term knowledge-based familiarity indicates, familiarity with a piece of technology relies on knowledge and information about it. According to Gefen et al. (2003), knowledge is built by gaining experience. From a technology's perspective, feedback helps a user to gain knowledge about the technology and to reduce confusing interaction by providing appropriate information. Feedback can be designed in different ways. When designing a feedback system for providing appropriate information to increase the user's familiarity with the technology, the designer has to answer the question, which information is useful for this purpose. In the context of affective technology, Heger et al. (2016) propose that "the system should constantly make visible to the user on which basis – that is which emotional state of the user the system has detected – the affective technology reacts" (p. 11), because the behavior of the affective system depends on precisely this. Besides, additional information on the behavior of the system can be useful feedback, since the ability to understand how an affective technology reacts after it has recognized a human emotion is significantly important (Heger et al., 2016).

> H5a: Feedback from an affective technology which provides information on the emotional state of the user increases their knowledge-based familiarity with the technology.

> H5b: Feedback from an affective technology which provides information on the system behavior increases a user's knowledge-based familiarity with the technology.

> H5c: Providing information on the emotional state of the user and the system behavior increases a user's knowledge-based familiarity with the technology more than if one of the two pieces of information is left out.

Besides information on the emotional state and system behavior, additional explanatory information on why a certain emotional state has been detected by the system and why the system reacts in a certain way can further help to become familiar with the system. The reason for this assumption is that familiarity relies on the predictability and anticipation of the other's behavior (Lewicki & Bunker, 1995). Familiarity means to develop an understanding of what is happening in the present (Luhmann, 1979). Thus, explaining to a user for which reasons an affective technology functions and behaves in a certain way will help them to anticipate future behavior.

> H5d: Feedback from an affective technology which explains why a certain emotional state of the user has been detected increases a user's knowledge-based familiarity with the technology.

> H5e: Feedback from an affective technology which explains why it behaves in a certain way increases a user's knowledge-based familiarity with the technology.

7.4.4 Method

Design and Participants. To test our hypotheses, we plan to conduct two experiments that are explained below. For the experiments, we chose an affective learning system that teaches the competence of writing a thesis, because many people and students in particular are familiar with e-learning environments and emotions are especially crucial for learning. For the purpose of simplification, we decided to focus the experiment on the recognition of only one emotion. As a low level of confusion is a learning-relevant emotion that, in contrast to frustration or boredom, is potentially beneficial for learning (D'Mello et al., 2014) and therefore should not negatively interfere with the perceived usefulness of the affective technology overall, we selected confusion. We did not select engagement as emotion of interest because, to the best of our knowledge, a reliable induction for this learning-relevant emotion does not exist. We specify the sample size to be 30 participants per group. Thus, we require a sample of 120 participants per experiment. We plan to recruit students from the local university for the experiments because scientific writing is a topic that is actually relevant for students.

Experiment 1: In the first experiment, our objective is to test the effects of different types of feedback content on affective technology acceptance (H1, H2, H3, H4, H5a, H5b, and H5c). Therefore, we want to conduct a 2 (emotion-related feedback vs. control) x 2 (system behavior feedback vs. control) between-subjects design.

Procedure. When the participant enters the laboratory, they are greeted by the experimenter who explains what affective technologies are and that we want to test a new form of affective learning system, which teaches the competence of writing a thesis. Afterwards, the experimenter leads the participant to the computer workplace, explains that the session is recorded by a camera through which the affects are recognized, and leaves the room. On the computer screen, the participant can read a short introduction that summarizes how a thesis should be written and structured. On the next page, we use a confusion induction (adapted from Lehman et al. (2013)) to bring participants into a state of confusion which could happen in a normal learning situation. For the confusion induction, two digital agents – a digital tutor (a professor) and a digital student represented with avatars – argue in a chat about the structure of a thesis. In contrast to what was written on the previous page, the digital tutor claims that, in the result's section, you should discuss in detail why unexpected results have occurred. The student, on the other hand, argues that the unexpected results should be explained in detail not until the discussion section. After the participant is asked about their opinion, one of the agents still disagrees. Then, the feedback manipulations (design options, see Table 7.14) are introduced through a pop-up window. Regardless of feedback manipulation, all participants are asked to explain their opinion to the two agents. Finally, participants complete the scales with behavioral intention, familiarity, trust, emotional self-reflexivity, manipulation checks, sociodemographic variables, and control variables on the computer and are debriefed by the experimenter.

Design options. The manipulation of the design options is presented in a pop-up window after the confusion induction is conducted. The design options are realized through different textual statements (see Table 7.14). In a paragraph under the statement of the feedback manipulation, the following text is presented in all conditions: “Please explain the arguments for your opinion”. Thus, the pop-up window in the control condition only differs in the absence of the first paragraph from the other conditions.

Emotion-related feedback	“The system has detected that you are confused.”
System behavior feedback	“The system has detected that explaining your opinion is the best learning strategy now.”
Emotion-related and system behavior feedback	“The system has detected that you are confused. The system has detected that explaining your opinion is the best learning strategy now.”
Control	*No text*

Table 7.14: Design Options in the Different Conditions of Feedback Content (Paper 12)

Manipulation Checks. We developed three manipulation checks each for emotion-related feedback and system behavior feedback manipulation. The items for emotion-related feedback focused on whether the participants noticed that the system told them about their emotional state. The items for system behavior feedback asked whether they noticed that the system told them a change in behavior.

Dependent and control Measures. The dependent and control measures will be measured on a 7-point Likert scale. The dependent measures are presented in the appendix. We adapt additional control measures for perceived ease-of-use and perceived usefulness from Gefen et al. (2003).

Experiment 2: In the second experiment, our objective is to test the effects of different types of feedback explanation on technology acceptance (H1, H2, H3, H4, H5d, and H5e). Thus, we will conduct a 2 (emotion-related feedback explanation vs. control) x 2 (system behavior feedback explanation vs. control) between-subjects design.

Procedure. The materials and procedure are mostly identical to the first experiment. In contrast to the first experiment, the design options that are prompted through a pop-up window are different and consist of the textual statements presented in Table 7.15.

Emotion-related explanation	"The system has detected that you are confused because you frowned. The system has detected that explaining your opinion is the best learning strategy now."
System behavior explanation	"The system has detected that you are confused. The system has detected that explaining your opinion is the best learning strategy now because this reduces your confusion."
Emotion-related and System behavior explanation	"The system has detected that you are confused because you frowned. The system has detected that explaining your opinion is the best learning strategy now because this reduces your confusion."
Control	"The system has detected that you are confused. The system has detected that explaining your opinion is the best learning strategy now."

Table 7.15: Design Options in the Different Conditions of Feedback Explanation (Paper 12)

Manipulation Checks. We developed three manipulation checks each for emotion-related explanation and system behavior explanation. The items for emotion-related explanation asked whether the participants knew how exactly the system identified the emotions they had. The manipulation check items for system behavior feedback explanation asked participants whether the system informed them why it reacted the way it reacted.

Dependent and Control Measures. Same as in experiment 1.

Data Analysis: We will use two-way ANOVA to test the main and interaction effects of the design options on familiarity. For testing all proposed hypothesis in unison, we will use PLS path modeling.

7.4.5 Discussion and Outlook

Affective technologies have great potential to serve humans; however, as they still are relatively new technologies, the acceptance of affective technologies is confronted with unique problems of understanding the functioning and behavior of the system that have only sparsely been addressed so far. With our research, we want to close this gap by examining design options that enhance the familiarity with the technology. Thus, we contribute to the literature by building a design theory for the context of affective technology with focus on the relation between feedback and familiarity.

Furthermore, we contribute to design theorizing in providing a methodological approach to test more than two design options of a design theory using a stringent methodological approach. Using two experiments that build on each other has several advantages. First, as the proposed hypotheses are tested two times with two independent samples, the theory is strengthened when it is not falsified in one or both experiments. Second, when we test the three design options (emotion-related feedback, system behavior feedback, and feedback explanation) in two separate experiments, we maintain a stringent experiment in which we vary as less information as possible in the different conditions. Finally, using two separate experiments offers the option of modification in the second experiment if the hypotheses cannot be supported. In this case, the theory can be adjusted for the next experiment without the need to recruit a large sample of participants for one unified experiment.

However, our research approach has limitations. As we want to use a pop-up window to deliver the feedback to the participants, a frequent interruption through pop-ups could be irritating in a real learning setting. Thus, future work could test how the feedback should be presented with regard to interruptions (e.g., testing the continuous presentation of the emotion-related feedback in a small part of the screen). Another limitation consists of the uncertainty whether all relevant hypothesized relationships can be supported in experiment 1 which could lead to the necessity of adapting experiment 2.

Appendix

Behavioral Intention to Use Affective Technology (adapted from Gefen et al., 2003 and Gefen, 2000)
I would use an affective learning system. I would allow an affective learning system to respond on the basis of my emotions. When having the choice between the same learning system with or without emotion recognition, I would rather use the one with emotion recognition.
Trust in affective technology (adapted from Gefen, 2000)
I believe that the affective learning system is trustworthy I trust the affective learning system. I'd trust the affective learning system to do the job right
Knowledge-based familiarity (adapted from Gefen et al., 2003 and Gefen, 2000)
I am familiar with the behavior of an affective learning system by using it. I am familiar with how an affective learning system responds. I can predict how an affective learning system responds. I understand why an affective learning system responds in the way it responds.

An affective learning system responds in the way I expect. An affective learning system responds reliably.
Emotional self-reflexivity – clarity (translated from Berking and Znoj, 2008)
Last week I could have stated clearly how I was feeling. Last week I was clear about what emotions I was experiencing. Last week I knew well how I was feeling.
Emotional self-reflexivity – comprehension (translated from Berking and Znoj, 2008)
Last week I was aware of why I felt the way I felt. Last week I understood my emotional reactions. Last week I knew what my feelings meant.
Emotional self-reflexivity – awareness (translated from Berking and Znoj, 2008
Last week I paid attention to my feelings. Last week I was aware of my feelings. Last week I dealt with my feelings.
Emotional self-reflexivity - sensations (translated from Berking and Znoj, 2008)
Last week my physical sensations were a good indication of how I was feeling. Last week I was physically well aware of my feelings. Last week I clearly realized when my body reacted noticeably to emotionally meaningful situations.
Manipulation checks – emotion-related feedback (self-developed)
The system told me what I felt. It was transparent to me which emotions the system recognized. The system notified me when it recognized a change in my emotions.
Manipulation checks – system behavior feedback (self-developed)
The system told me how it reacted. It was transparent to me how the system reacted. The system notified me when it changed the learning strategy.
Manipulation checks – emotional-related feedback explanation (self-developed)
The system told me how it recognized my emotions. It was transparent to me on which basis the system recognized my emotions. The system notified me when my bodily reactions changed.
Manipulation checks – system behavior feedback explanation (self-developed)
The system told me why it reacted the way it reacted. It was transparent to me why the system reacted the way it reacted. The system notified me why it changed the learning strategy.

Table 7.16: Measurement Items (Paper 12)

8 Track 4: Ethical Implications

8.1 Paper 13: Making Use of Facebook Comments for Upstream Engagement

Title	**Making Use of Facebook Comments for Upstream Engagement**
Authors	Katharina Jahn katharina.jahn@uni-siegen.de Oliver Heger oliver.heger@uni-siegen.de Marius Müller marius.mueller@uni-siegen.de Bjoern Niehaves bjoern.niehaves@uni-siegen.de
Publication Type	Conference Proceedings
Publication Outlet	International Conference on the Ethical and Social Impacts of ICT
Outlet Information	–
Status	published
Full Citation	Heger, O., Jahn, K., Mueller, M., & Niehaves, B. (2017). Making use of facebook comments for upstream engagement: A systematic approach. *CEPE/ETHICOMP 2017. Turin, Italy*

Table 8.1: Fact Sheet Paper 13

Making Use of Facebook Comments for Upstream Engagement

Abstract

Deliberative activities constitute an essential part of Responsible Research and Innovation (RRI). Within the deliberative dimension of RRI, so-called "upstream engagement" covers activities which try to legitimize, authorize and prioritize research agendas and intentions. Although upstream engagement is an effective approach to include the public in research, its implementation requires time and effort. To address this challenge, we have developed a systematic, Facebook-specific approach for capturing feelings, ideas, options and priorities towards a certain type of technology and integrating them into technology-related research. To illustrate our proposed approach, we have applied it to 'virtual reality' and 'affective technology' as exemplary cases.

Keywords— Responsible research and innovation, upstream engagement, technology acceptance, design, social media, Facebook

8.1.1 Introduction

Deliberative activities constitute an essential part of Responsible Research and Innovation (RRI). They aim at involving a broad range of perspectives from stakeholders into the innovation process of technology design (Bessant, 2013). Deliberation includes activities whose purpose is to optimize decision processes and improve socio-technical outputs regarding the emergence of new technologies (Fisher et al., 2006). By including multiple perspectives, the outcome of the RRI process is expected to be more socially desirable.

Within the deliberative dimension of RRI, so-called "upstream engagement" covers activities which try to legitimize, authorize and prioritize research agendas and intentions (Jackson et al., 2005). Wilsdon et al. (2005) state that upstream engagement „encourages dialogue between scientists and the public to move beyond competing propositions, to a richer discussion of visions and ends", which leads to a broader consensus on the technology at hand. With the objective of ensuring a broad consensus, implementing upstream engagement in a design-oriented innovation process, firstly, means to find out what determines the social acceptance of the technology to be designed. Secondly, implications are then to be derived which inform the design of the technology.

Although upstream engagement is an effective approach to include the public in research and, in doing so, to implement the deliberative dimension of RRI, its implementation requires time and effort. This challenge needs to be addressed by the executing researchers. For instance, setting up workshops or conducting interviews can be time consuming for the researchers and the interviewees as well. To address this challenge, the paper at hand provides a way for a quick and easy implementation of upstream engagement. The first research question here is:

Research Question 1. How can upstream engagement be implemented quickly and easily?

Moreover, many people are not interested in participating in research activities when you ask them to. According to a poll (Shah & Castell, 2011), most of the respondents do not want to get personally involved. Facebook being one of the most widespread social networks offering plenty of public dialogues provides data that can be made use of to capture a wide range of perspectives in the public. Thus, the paper is led by a second research question:

Research Question 2. How can already existing Facebook dialogues be utilized for upstream engagement?

Following both research questions, we have developed a systematic, Facebook-specific approach for capturing feelings, ideas, options and priorities towards a certain type of technology and integrating them into technology-related research. The approach is based on the 'typical' procedure of sentiment detection (Liu, 2010) and methods from netnography (Kozinets, 2010). To illustrate our proposed approach, we have applied it to 'virtual reality' (VR, 360° virtual environment with headset) and 'affective technology' (AT, technology that detects the emotional state of its user) as exemplary cases. Finally, we discuss implications for future innovation (processes) of technology design as well as limitations and outlook of our work.

8.1.2 Literature

8.1.2.1 The Role of Upstream Activities in RRI

Modern perspectives in social sciences consider public social actors as valuable contributors in scientific discussions and discourses. Scientists are moving away from treating the public as a passive actor that is obliged to deal with the results or products science delivers (Kouper, 2010). Hence, by deliberately engaging these actors, a broader spectrum of societal aspects regarding innovative research can be taken into account when it comes to designing ICT (information and communications technology). Efforts to actively address the public are increasingly encouraged by the ongoing advancement of modern communication technologies (e.g. social media) "which enable active co-operation of user communities in co-creation and diffusion" (Bessant, 2013). User engagement, thus, can have a significant impact on the adoption of a technological innovation (von Hippel, 2009).

In order to manage public engagement, the so-called upstream is of major interest. By involving the public, development processes of emerging and innovative technologies can be formed collectively (Jackson et al., 2005). This collective knowledge is needed to address the technology's future impact, which „lies in the future and so in the emergence phase is still somewhat uncertain and ambiguous" (Rotolo et al., 2015). The ongoing dialogue between researchers, decision makers and public entities is of an interactive and mutual nature and leads to rich and deep data, whereas polls only "measure 'top of the head' public views" (Warburton et al., 2008).

8.1.2.2 Challenges for Upstream Engagement

The engagement of the public comes with challenges and necessities that need to be addressed by the executing researchers. On the one hand, upstream engagement is affiliated with a fairly high amount of effort when it comes to actively including the public. Opportunities for stepping into a dialogue need to be granted, e.g. by setting up workshops or conducting interviews, which can be time consuming for the researchers and the interviewees as well (King, 1998). On the other hand, it can be hard to encourage the public in actively engaging in research. According to a poll conducted by Shah and Castell (2011), 56% of the respondents do not feel informed about science and research, although a majority of them think positively of science. Furthermore, "most do not want to be personally involved". In addition to that, a certain extent of ignorance and lack of knowledge is apparent which can lead to severe tensions between the public and scientific domain, calling for academic enlightenment (Skyes & Macnaghten, 2013).

In 2014, an event organized by the National Co-ordinating Centre for Public Engagement (NCCPE) revealed several challenges the domain of public engagement will encounter in the near future (NCCPE, 2015). For instance, priorities of public social actors are constantly changing, which implies that methodological approaches need to be flexible and, thus, more efficient. Additionally, NCCPE suggests that researchers should partly move away from obvious participants (e.g. social actors that are exposed to a certain technology in their working domain). Instead, the engagement of a broad diversity covering varying and distinct opinions seems promising.

The challenges mentioned above are in line with our systematic approach, reducing the effort of data gathering and analysis while involving large online communities that voluntarily share their opinions with others.

8.1.2.3 The Potential of Social Media Upstream Engagement

Social media has potential for engaging the public in research as it attracts a huge amount of different users who are exchanging opinions on all sorts of themes. Therefore, upstream engagement can benefit from analyzing the data they provide. Netnography, a methodological approach that applies ethnographic methods to the context of online communities (Kozinets, 2002), provides a research methodology that can help to analyze such data. In the following, we explain netnography in the context of social media and explain how this approach is relevant for upstream engagement.

8.1.2.4 Netnography

Netnography is “a specialized form of ethnography adapted to the unique computer-mediated contingencies of today’s social worlds” (Kozinets, 2010). Netnography provides a five step approach that considers the specific requirements of online research (Kozinets, 2010): First, the research questions are defined. Second, the community where the research should be applied is identified. In the following third step, data collection is carried out under ethical considerations. In the fourth step, the data is analyzed. Finally, the results of the research are presented to the public.

8.1.2.5 Sentiment detection

In this research paper, we develop an approach to specify the data collection and analysis steps of netnography for upstream engagement in applying sentiment detection. Sentiment detection, as a form of text mining for identifying opinions and emotions towards entities in different sorts of texts (Liu, 2010), provides useful methods for analyzing comments in social media networks and has been successfully applied on various social media platforms such as Facebook (Ortigosa et al., 2014) or Twitter (Kouloumpis et al., 2011). Thus, sentiment detection can be considered as an appropriate method to analyze social media content for upstream engagement.

8.1.3 Development of a Systematic Approach

Upstream engagement can be used to explore people's feelings and develop "ideas, options and priorities" with the public (Warburton et al., 2008). For the purpose of including the feelings, ideas, options and priorities towards a certain type of technology into technology-related research, this paper proposes the following systematic approach for deriving design implications for a technology from its determinants for acceptance and rejection. By 'acceptance' we mean a positive attitude towards a technology, by 'rejection' a negative attitude. The approach follows the 'typical' procedure of sentiment detection, but without focusing on its technical/algorithmic implementation.

Step 1 – Selecting suitable Facebook posts for a type of technology in one or more relevant application fields: Choose posts on Facebook which provide appropriate content and comments. Select such posts which address relevant stakeholders, have a high number of comments, are published by reputable platforms and are rather neutral and not polemical.

Step 2 – Categorizing comments into acceptance, rejection, and neutral stance/no opinion: Collect all comments of the posts you have selected and assign them in whole or in part to one of the three categories.

Step 3 – Extracting the reasons for acceptance and rejection: Go through all statements of both the acceptance and rejection category and, if any, extract the reason mentioned for accepting or rejecting the type of technology.

Step 4 – Clustering the extracted reasons: Go through all reasons you have extracted and build clusters.

Step 5 – Identifying determinants for acceptance and rejection: Determine the core of the clusters you have built. For the designation of the clusters, we recommend to take acceptance theories from the field of Information Systems into account, which contain 'classical determinants' such as "ease of use" or "usefulness".

Step 6 – Deriving design implications: Find design requirements which meet the determinants for acceptance of the technology and oppose the determinants for rejection.

An example to illustrate the approach is given in Table 8.2:

Step 1: Post by 'Cambridge English': "Some people say that virtual reality will change education. What do you think? How do you think virtual reality will change language learning?" (22 comments)

Step 2: Comment: "I believe virtual learning resources are in and are rather helpful (-> acceptance) but nothing can replace a good teacher. I have studied different languages and I have used virtual studying devices just as a mere complement but always have turned to my teacher (-> rejection)".

Step 3: Reason for acceptance: "virtual learning resources are in"; "are rather helpful"; reason for rejection: "nothing can replace a good teacher, [...] always have turned to my teacher"

Step 4 & 5: Determinant for acceptance: "current trend", "usefulness"; Determinant for rejection: "Replacement of teacher"

Step 6: Design a useful virtual reality tool as a complement to existing teaching formats.

Table 8.2: Example „Virtual Reality for Education" (Paper 13).

8.1.3.1 Ethical issues of our approach

When doing netnography, ethical issues have to be considered regarding data collection. For example, when collecting information from closed groups in online communities, it is unclear whether participants agree that their posts are used by researchers and are made available publicly. We solved this issue in this paper by only collecting data Facebook posts which are publicly available. However, when adopting our approach for other cases, ethical issues should be considered, e. g. through obtaining the consent of participants in closed communities.

8.1.4 The Exemplary Cases of 'Virtual Reality' and 'Affective Technology'

To illustrate our proposed approach, we have applied it to 'virtual reality' (VR) and 'affective technology' (AT). By VR we mean a 360° virtual environment generated by a headset. VR technology is widely known due to reports on 'Oculus Rift' or 'HTC Vive'. In contrast to VR, AT has not yet established itself in the consumer market and remains rather unknown to the general public. ATs are systems which can sense and/or generate human emotions (e.g. happiness, anger, fear). An affective learning system, for instance, could identify boredom or frustration and regulate its level of difficulty accordingly. The entire data has been "manually" coded by the first two authors of this paper.

Step 1: For VR, we chose two application fields (game and health care) and three posts per application field. For each post, we coded up to 50 comments (in total 150 comments for game and 139 comments for health care). For AT, we selected three posts on emotion recognition technology across different application fields with 144 comments in total. The main challenge of this step was to find appropriate posts which were not too polemical or political and encouraged multiple users to provide reasonable comments. Serious news organizations, such as CNN or the German news broadcast "Tagesschau", proved to be the most promising providers of such posts, as they tend to report neutrally and have a high number of subscribers.

Step 2: In the second step, we categorized the comments into one or more of the three categories "acceptance", "rejection" or "neutral". As can be seen from Table 8.3, applying VR in the game context received the highest amount of acceptance and the lowest amount of rejection, followed by applying VR for health care. In contrast, the commentators are more cautious regarding AT which received less than half the amount of acceptance than the other two application fields and the highest amount of rejection. Notably, there was a high amount of neutral comments in all application fields, ranging from 71 neutral comments for VR for health care to 86 neutral comments for VR in the gaming context.

	acceptance	rejection	neutral
VR/game*	40	28	86
VR/health care	38	31	71
AT	18	41	85

*> 150 because a comment can partly express acceptance and rejection

Table 8.3: Results of Step 2 "Categorizing Comments" (Paper 13).

Categorizing comments into acceptance and rejection requires a high amount of interpretation. Especially (supposedly) ironical comments were subject of discussions. If the first two authors could not agree on a comment, the third author decided. However, the vast majority of comments could be clearly categorized.

Step 3: In the third step, we extracted the reasons the commentators had for accepting or rejecting the respective technology. For VR in the gaming context, reasons for acceptance were brought up 38 times. In contrast, there were only 17 times that commentators mentioned reasons for accepting AT. When looking at reasons for accepting, the reverse picture emerges: For AT, reasons for rejection were mentioned 40 times, whereas they were only mentioned 29 times for each VR context. The results of step 3 are presented in Table 8.4.

	reason for acceptance	reasons for rejection
VR/game*	38	29
VR/health care	21	29
AT	17	40

*< 40 reasons for acceptance because an acceptance comment does not necessarily give a reason. But in contrast, a comment can give more than one reason (cf. rejection).]

Table 8.4: Results of Step 3 "Extracting Reasons" (Paper 13).

Similar to step 3, identifying the reason for acceptance or rejection requires interpretation. To avoid losing valuable data, we decided to follow a rather generous interpretative approach. All sentences or terms which could potentially provide information on why a user expresses a certain opinion have been extracted.

Step 4 & 5: Since step 4 and 5 are strongly interrelated and are done iteratively, we have summarized both steps to one. We started by clustering all reasons we have collected and abstracting them into collective terms, which resulted in a list of determinants for acceptance and rejection per application field (shown in Table 8.5).

Certainly, the clusters and terms were not always clear-cut. That is why the first two authors had to discuss and find common solutions. Since the objective of this step is to find determinants for acceptance (and rejection), acceptance theories from the field of Information Systems could help to build clusters and find collective terms. The technology acceptance model (TAM) (Davis et al., 1989) is a well-known example for an acceptance theory, which, for instance, could provide us the determinant of 'usefulness'. However, these theories could not provide us any application field-specific determinants, which, in our case, are grounded in the data.

Step 6: In the final step, we derived design implications building upon the identified determinants from the previous steps. Table 8.6 exemplarily shows one design implication addressing selected determinants for acceptance and rejection for each of the application fields. For example, in the case of VR in the gaming context, a determinant for acceptance consisted of the possibility to experience phantasies in virtual reality games. Thus, a design implication that might be beneficial for a certain stakeholder ("fans of novels") is using VR to bring phantasies alive (e. g. through adopting novels

	determinants for acceptance	determinants for rejection
VR/game	future-orientation, usefulness, fun, curiosity, past experience, experiencing phantasies, physical activity, health	motion sickness, third party's unfamiliarity, distance from reality, mental-related changes, social-related changes, ugly design, lacking maturity, distraction from other activities, fear of addiction
VR/health care	physical enhancement, mental enhancement, usefulness, curiosity, future orientation, familiarity	lacking usefulness, lacking understanding of functionality, lacking trust in media/functionality/ doctor, strain, technology hostility, motion sickness, addiction, high costs, time, fear, potential harm
AT	trust in functionality/ government/structural assurances, low costs, usefulness, just world beliefs, health, curiosity, social-related changes, physical enhancement, mental enhancement, technological advancement	lacking trust in functionality/government/manufacturer, lacking usefulness, high costs, surveillance, "sick" behavior, fear of addiction, cheaper substitute, technology hostility

Table 8.5: Results of Step 4 & 5 "Clustering & Identifying Determinants" (Paper 13).

in a VR game). In contrast, a determinant for rejecting VR was called "expected social changes related to isolation". To address this issue, VR games could be designed in a way that encourages – or forces – people to meet in reality in order to be successful in the game. In the context of AT, one determinant for accepting this technology was trust in structural assurances. A resulting design implication to foster acceptance could therefore be structural assurances in form of external safety structures (e. g. through laws that regulate the use of AT for companies). For commentators who rejected AT, surveillance was an important issue. Thus, a promising design implication might be to implement privacy in design through transparency and data avoidance.

Step 6, the translation of determinants for acceptance and rejection into design implications, highly relies on the technical and conceptual expertise as well as intuition of the persons conducting this step. However, we believe it to be important to include this step, since the main objective of RRI and upstream engagement is to make a signifi-

	determinant	design implication
VR/game	+experiencing phantasies	Use VR to bring phantasies alive (e.g. based on novels with a strong fan base)
	-social-related changes (isolation)	Meeting people in reality can be designed as a part of the game.
VR/health care	+physical enhancement	Use VR to teach doctors how to move correctly.
	-lacking understanding of/trust in functionality	The tool should provide sophisticated help and be designed "transparently" (what is it doing? which data does it collect? which dangers exist? etc.).
AT	+trust in structural assurances	Provide the users of AT with external safety structures (e.g. regulations, laws, guarantees etc.).
	-surveillance	Do not forward information about the user or, if necessary, make transparent who gets which information.

Table 8.6: Results of Step 6 "Deriving Design Implications" (Paper 13).

cant impact on the outcome of research and innovation. Thus, the design implications should reflect for which purposes a technology is to be designed and which features and functions it should have.

8.1.5 Discussion

In this paper, we have proposed a systematic approach that makes use of Facebook comments for upstream engagement and applied it to three exemplary cases (VR for gaming, VR for health, and AT). Applying our method on the exemplary cases suggests that Facebook comments can be used to supplement upstream engagement, even though some hurdles still exist. Especially if a researcher of a design-oriented research project is interested in identifying feelings, ideas, options, and priorities from the public before having built a prototype, the approach proposed in this paper can be of help.

Considering the challenges of upstream engagement when integrating public actors into research activities, our developed approach exhibits several advantages compared to the ways upstream engagement is usually deployed. First, our approach comes with a tremendously reduced amount of effort. The researcher does not need to actively step into a dialogue with participants, lessening the methodological burden (e.g. the organization of interview appointments). Hence, the approach requires less temporal and monetary resources, representing economic advantages over other methods. Second, potential research subjects (i.e. public actors) do not need to be encouraged to unveil their opinions and thoughts on the topic of interest. They post them voluntarily and publicly via social media platforms and, thus, are intrinsically motivated to do so, requiring no further external incentives. This comes with the advantage that researchers can access a broad variety of opinions and discussions, which leads to multifaceted perspectives enriching the researcher's often narrow or very specific point of view. Additionally, the aforementioned tensions between the public and scientific domain can be (partly) relieved by getting closer to the people's real-world context. Due to the flexible and efficient nature of our approach, it can be deployed at a very early stage of scientific projects. Thus, it allows quick adaptations and modifications of research activities and agendas. In addition to that, knowledge gathered via the approach at hand is able to serve as a "ground truth" researchers can build upon, legitimizing future initiatives.

However, our approach is not without limitations. First, the representativeness of social media users and especially commentators can be questioned. Nevertheless, this depends on the researchers' target group. Second, upstream engagement requires an intensive dialogue between research and public which the proposed method can only partly address. For example, a technology that is too new to be discussed in online communities could not be examined with our approach. However, most technologies are discussed rather early within relevant online communities, as the exemplary case of AT shows. Lastly, there are several comments that are not substantial or contain irony which might complicate the correct categorization of comments and extraction of reasons. Thus, one should be aware that the method of sentiment detection relies to some degree on subjectivity.

A final aspect that should be considered are ethical implications of the developed approach. It is important to note that, regarding ethical concerns such as privacy, the researcher has the responsibility to take care of ethical issues that might emerge from analyzing posts from closed communities or communities in which the commentators do not expect to be cited from. Additionally, stemming from the issue of a possible low representativeness from comments and social media sites, it is possible that not all stakeholders are taken into account through analyzing the comments. Thus, researchers should bear in mind that design implications resulting from this approach might be incomplete and that there might be stakeholders that should be addressed differently.

In conclusion, using Facebook is a promising approach for researchers who plan to apply upstream engagement and Responsible Innovation. Furthermore, our approach is especially useful in a very early stage of product development. This offers the opportunity to automate the approach with technical solutions from sentiment detection.

8.1.6 Acknowledgements

This article was supported by the projects INEMAS (No. 16SV7235) and ELISE (No. 16SV7512) funded by the German Federal Ministry for Education and Research (BMBF).

8.2 Paper 14: More Than Ticking Off a Checklist?

Title	**More than Ticking Off a Checklist? Towards an Approach for Quantifying the Effectiveness of Responsible Innovation in the Design Process**
Authors	Katharina Jahn katharina.jahn@uni-siegen.de Hendrik Kempt hendrik.kempt@uni-siegen.de Tanja Joan Eiler tanja.eiler@uni-siegen.de Oliver Heger oliver.heger@uni-siegen.de Alla Machulska alla.machulska@uni-siegen.de Armin Gruenewald armin.gruenewald@uni-siegen.de Tim Klucken tim.klucken@uni-siegen.de Rainer Brueck rainer.brueck@uni-siegen.de Carl Friedrich Gethmann carl.gethmann@uni-siegen.de Bjoern Niehaves niehaves.bjoern@uni-siegen.de
Publication Type	Conference Proceedings
Publication Outlet	Workshop “Ethik und Moral in der Wirtschaftsinformatik”
Outlet Information	–
Status	published
Full Citation	Jahn, K., Kempt, H., Eiler, T. J., Heger, O., Gruenewald, A., Mach, A., Klucken, T., Gethmann, C. F., Brueck, R., & Niehaves, B. (2020). More than ticking off a checklist? Towards an approach for quantifying the effectiveness of responsible innovation in the design process. *Workshop Ethik und Moral in der Wirtschaftsinformatik 2020, Potsdam, Germany*

Table 8.7: Fact Sheet Paper 14

More than Ticking Off a Checklist? Towards an Approach for Quantifying the Effectiveness of Responsible Innovation in the Design Process

Abstract

Recent technological developments highlight the need to address responsible technology designs. One of the best-known approaches to integrate ethical design into artifacts is Value Sensitive Design (ValSD), which relies on the involvement of stakeholders to identify values important for different groups. However, to which degree ValSD and related approaches actually change the outcome of technology development instead of being a mere legitimization for addressing ethics is still unclear. Thus, we aim to investigate how the transparency about the effectiveness of responsible innovation in the design process can be improved. To address this, we propose an approach that is based on the rules of ethical and moral discourse and aims to identify norms acceptable to moral agents. To increase transparency, we differentiate between the initial design solution in a project, the solution identified as feasible after ethics workshops, and a "regulative idea", describing the best solution under conditions of unlimited resources.

Keywords— responsible innovation, virtue-sensitive design, value-sensitive design, rational agent-centered design

8.2.1 Introduction

As mobile technologies, virtual reality, and even artificial intelligence diffuse into our daily life, a range of different tasks gets easier or more enjoyable. To only name two application areas, mobile technologies can facilitate health behavior (Bojd et al., 2018), and virtual reality can enhance cognitive performance (Banakou et al., 2018). In line with this, technologies attain a new level of complexity, and their impact on our life becomes more and more multifarious, which highlights the need to address the question of responsibility for researchers. For example, in the context of health, potential users can be concerned with different aspects of the design, which designers might not be automatically aware of, including empathy, privacy, and diversity (Dadgar & Joshi, 2018; Mueller & Heger, 2018). Therefore, it becomes evident that considering ethics as a mere byproduct is not sufficient. Rather, ethical considerations ought to be an integral part of the complete design process, with ethical issues not only accompanying the process but fostering change in the designed artifact in a way that benefits individuals and society.

Different methods already exist to account for responsibility in the design process serving as a framework for responsible innovation. The probably best-known approach, value sensitive design (ValSD, Friedman & Hendry, 2019), provides a value-oriented method in which the researchers involve stakeholders to investigate values and integrate a solution for possible value conflicts into the artifact design. On the other hand, a method that is more grounded in philosophical theory and relies on the individual researcher, virtue sensitive design (VirtSD, Vallor, 2016), proposes that design researchers can account for responsibility through their own education. Although additional approaches exist, such as Participatory Design in which the participation of users is specifically highlighted (Iversen et al., 2010; Spinuzzi, 2005), we focus on ValSD and its virtue-related counterpart VirtSD because they provide adequate examples to illustrate potentials for integrating ethical design into the design artifact.

Research Question. How can we examine and measure the effectiveness of responsible innovation in the design process?

To address these issues, we develop a method that 1) increases transparency in the design process with regard to which ethical goals are abandoned or maintained in the design process, 2) provides a measure of how effective the design process was in being ethical, and, as a consequence, 3) can contribute to motivate researchers and designers to implement responsibility into IT artifact design more rigorously.

Our paper is structured as follows. First, we present ValSD, VirtSD, and their respective shortcomings and sketch a theory that can avoid those issues. Next, we provide the case of developing a virtual reality technology for smoking cessation in which we apply our approach. In the following section, we illustrate the method for our case that contributes to answering our research question by making transparent in which areas an ideal solution in the form of a regulative idea could, and could not, be integrated and explain how this helps with comparing design solutions. Finally, we summarize the contributions of our approach and the next steps to develop it further.

8.2.2 Approaches to Ethical Design

8.2.2.1 Responsible Innovation

Research or innovation projects have two dimensions that require ethical reflection. Both their artifact (usually a particular prototype or demonstrator), as well as the method of designing said artifact (e.g., experiments), ought to be considered when discussing the normative requirements of any research project. We regard both those dimensions as substantial ethical issues and hold that ethically convincing responsible research and innovation (RRI) ought to be sensitive to both.

Turning to the available theories on RRI, we found many approaches to be helpful for instructions on how to identify and approach ethical issues in research and innovation but also concluded that they were ultimately lacking. In the following, we will shortly discuss the benefits and shortcomings of two of the main approaches that motivated and necessitated a new path taken on for this project.

8.2.2.2 Value-Sensitive Design

Value-sensitive design (ValSD, Friedman & Hendry, 2019; Friedman et al., 2017; van den Hoven, 2013) has become one of the most common approaches to RRI. The idea behind ValSD consists of developing an accessible, systematic evaluation method of design and innovation processes. Values act as the central evaluative measure and are usually understood as "referr[ing] to what a person or group of people consider important in life" (Friedman et al., 2008), without being committed to specific ethical theories. Due to the flexibility of the purported concept of "value", many areas of ethical concern can be covered, and in fact, some can be discovered using the terminology developed in ValSD. The concentration on direct and indirect stakeholders, values of both stakeholders and designers, as well as the integration of ethical evaluations in the design process, aims at the coverage of ethics issues from broad perspectives and provides a method that has led to a more ethically aware approach to RRI.

However, ValSD does have some problematic, underappreciated philosophical implications that we aim to avoid. Mainly, the concept of "value" allows for several different philosophical interpretations without ValSD providing a philosophically sufficient definition in its method (Albrechtslund, 2007; Davis & Nathan, 2015; Jacobs & Huldtgren, 2018). This variety in philosophically demanding value-concepts causes two main issues. For one, it remains unclear if all ValSD-projects share the meaning of their ethical core terminology. Two different concepts of "value" can lead to inconsistent results in the evaluation of a given design-problem, suggesting a certain relativity and thereby weakening the binding validity claims of applied ethical research. Second, without a shared definition of the value-concept, it increases the difficulty of implementing values in RRI to a desirable degree of impact and precision. Both the number of relevant "values" and their weight in ethical deliberations remain unclear in ValSD approaches, due to the opaqueness of the presupposed concept of value. The consequence consists of the risk of having ethical considerations being an afterthought or appear to be checked off a list without any substantial interaction between the way a research project is conducted and the values it is supposed to implement. Moreover, while we acknowledge that this is not how ValSD was conceived to work, we think it is crucial to avoid these exploitable weaknesses.

8.2.2.3 Virtue-Sensitive Design

On the other hand, a philosophically more elaborate position has been put forward by Shannon Vallor (Vallor, 2016). The so-called VirtSD aims to emancipate researchers in their ethical thinking. This is done by educating the empirical researchers and engineers in ethical thinking, rendering them "virtuous." The idea behind this approach is to guarantee ethical deliberations being an integral part of scientific research without ethicists coming in at a later stage and being perceived as a "correction" to the otherwise non-ethical research. Furthermore, while we recognize the need to make ethical deliberations more obvious to engineers and scientists, we hold that relying on their expertise alone is an undue burden and thereby increases the risk of failing to guide RRI processes. It can be such an undue burden because ethical considerations are not trivial and often require valuable resources of time and knowledge (Gethmann & Sander, 1999). Complex deliberations of possible ethically relevant consequences and familiarity with arguments from applied ethics are not merely guaranteed through a "virtuous scientist," but would require them to enter an active philosophical discourse – which can hardly be expected without having studied philosophical discourses extensively. Thereby, we advocate for an active role of ethical specialists in RRI, which evidently does not dismiss the idea that practitioners can contribute to ethical discourses.

8.2.2.4 Rational Agent-Centered Design

We claim that both approaches do provide insight and partially reliable guidance for RRI. However, with both methods, we identify two reasons requiring a new approach. The first reason consists in the difference of ethical questions arising from research and innovation processes on the one hand and the use of their results on the other hand. The assumption seems to be that "values", or a virtuous scientist, can both identify the procedural ethical requirements of a research project, as well as assess the results of said project. These methodological demands require a more elaborate ethical vocabulary than one concept that is expected to do the heavy conceptual lifting. The second reason is the lack of measurability and quantifiability of the effectiveness of the ethical considerations. The premise of a two-sided approach to ethical deliberations is motivating our approach. The first step requires deliberations on agents whose view on ethics issues in the system are relevant in a research project ("stakeholders"). Not everyone claiming to have a relevant say in ethical discourses does have such a claim. The challenge of RRI in this first tier is to identify those stakeholders and their claims in a given research project. In this context, arising ethical "issues" are merely those that can cause a conflict between two claims (similar to the concept of "value-conflict" in ValSD). We hold that ethical deliberations ought first to find those potential conflicts of claims and then to provide the discursive space to resolve those conflicts. Especially in human-centered research, those deliberations prove to be of extraordinary relevance due to several ethically established rules, such as the principle of non-instrumentalization (i.e., the "dignity" of human agents) that may conflict with a project design to maximize possible research results. Especially in psychological experiments, this principle can stand in conflict with success-requirements of the experiment, in which participants have to be tested on something they have to remain unaware of. For this, depth and extent of consent have to be taken into consideration as the primary guidance on to what extent personal rights can be temporarily infringed. Additionally, design-time issues may arise, such as designers' preferences regarding the content of the system leading to bias in the design. To address these issues, we refer to basic concepts developed in discourse-ethics, which is a well-established theory of normative ethics with some application in technology assessment. By using discourse ethics, we avoid the problem previously encountered in ValSD, where the lack of philosophical elaboration caused problematic methodological implications.

In basing our approach on the rules of ethical and moral discourses in the constructivist tradition (c.f. Grunwald, 1999; Kamlah & Lorenzen, 1967), we have to assume to deal with rational moral agents (c. f. Grunwald, 1999). Thereby, the results of our deliberations ought to be acceptable to any moral agent, rendering "acceptability" a key concept in our approach. This "rationalistic approach" produces norms that ought to be considered valid to any rational agent based on their insight into the discourse. We presuppose an instrumentalistic concept of rationality, as we reconstruct the adequacy of choosing specific means to realize one's purposes and preferences. Only a reasonable objection will re-trigger the discourse to come to a resolution of said objection. It also allows for a concept of responsibility. A particular research project design is only acceptable if the designers are in control of (i.e., can be made responsible for) any reasonably anticipated risks as well as taking precaution for their research results not to be misused in ill-intended applications. Due to the issues of "unknown unknowns", the requirements of "acceptability" are limited to discoursively determined risk-assessments. It thereby covers both the design process as well as possible results from the prototype/demonstrator.

In the following, some examples illustrate how this method produced operationalizable results in our research project and how we intend to develop it further.

8.2.3 Case Description

The project aims to support patients in the treatment of substance dependence diseases. In doing so, we rely on the dual-process model of addiction (Deutsch & Strack, 2006), which states that two qualitatively different mental process types regulate output (behavior) in relation to a certain input (stimuli): the impulsive and the reflective processes. While reflective processes (e.g., motivation, skills, knowledge) are consciously perceived, impulsive processes (e.g., approach, avoidance, attention) are automated. In a healthy individual, there is a balance between the two types, in the case of addiction; however, the impulsive processes become dominant, which usually leads to an emergent behavior towards addiction-related stimuli.

We use the Approach Avoidance Task (AAT) (Machulska et al., 2015; Machulska et al., 2016; Rinck & Becker, 2007) as a basis to restore the balance between these two processes. In the desktop version, users receive indirect instructions to which they should react using a joystick. For example, all addiction-related stimuli should be pushed away, and all neutral stimuli should be pulled. This has to be done as fast and accurately as possible in order to (re-)train automatic behavior. This procedure is called Cognitive Bias Modification (CBM), as attempts are made to retrain the approach bias.

One problem of CBM is the relatively high dropout rate (Beard et al., 2012; Machulska et al., 2016; Schoenmakers et al., 2010)). Therefore, our project aimed at developing a VR application, which should increase the motivation of users by design variables, such as high immersion, presence, and embodiment.

To transfer the AAT procedure into three-dimensional space, the users are located in an office room with a table in the middle. A cardboard box stands in front of the table, behind it a garbage can. The participants receive the indirect instruction to pull all stimuli outlined in blue into the cardboard box and to push all stimuli outlined in red into the garbage can. This preserves the vital arm movements of the AAT. In the process, the reaction times for the calculation of the approach bias are recorded.

Further levels, which can represent and train different everyday situations, the acquisition of medical data, as well as gamification elements (e.g., badges, leaderboards), are planned. If successful, the developed VR application will be included in the portfolio of a regional clinic.

8.2.4 Ethical Conflicts and Their Resolutions

Over the course of several ethics workshops, we have identified several norms and multiple possible ethical conflicts associated with the design goals of the project. In the following, we selected the norms a) privacy, b) exclusion of participants, and c) societal changes for which we describe selected ethical conflicts below to illustrate how our approach can be applied.

The first area of conflict relates to the privacy norm. In the VR context, data about the users' progress and performance can be saved. Thus, the question arises if it is saved, at which place it is saved and who decides which persons have access to the data. Second, we have identified ethical conflicts related to the possible exclusion of participants due to the used technology. Due to the high diffusion of smartphones in our daily life, almost everyone can use the smartphone application. However, especially for the VR application, cybersickness can exclude participants and lessen their motivation to use the training.

Finally, the norm of preventing societal changes not beneficial to all individuals may also be violated in the long term. When the application is so successful that almost everyone can stop smoking if they want to, individuals who intend to continue smoking might be faced with stigmatization from their environment.

8.2.5 Method

We have conducted four ethics workshops since the project's start, in which philosophers of technology, computer scientists, designers, information systems researchers, psychologists, and current smokers participated. The results of the ethics workshop related to selected examples of implementations of the above-mentioned ethical issues are listed below. Our method consists in differentiating between three types of design solutions related to the ethical issue. 1) The initial solution relates to the design solution that was proposed by the team without an ethics workshop. 2) The solution after the ethics workshop relates to the best feasible solution during the project time that we decided on after conducting the ethics workshop. 3) The regulative idea, a term we borrow from Immanuel Kant's work, refers to the solution that would best address the respective conflict if we had unlimited resources during the project. A regulative idea, in this regard, gives agents an idealized projection which serves as a horizon to which they can orient their actions. It is not required to fulfill the regulative idea in full to count as ethically justified.

8.2.5.1 Addressing the identified ethical conflicts with design

Regarding the privacy norm, the initial solution lies in saving the data of the users pseudonymized on the computer on which the VR application is running. However, one of the ethics workshops revealed that this solution might be insufficient because users may have concerns that the researcher or third parties can recreate the pseudonymized code. To address these concerns, we concluded that it was sensible to provide a higher degree of data security as default, including encryption, and give possible users of the validated, final VR application used for training purposes the opportunity to decide against data collection (which is, at the current stage, only needed for research purposes and not for training purposes).

Norm	Initial Solution	Solution after the Ethics Workshops	Regulative Idea
Privacy	saving data pseudonymized on VR-computer	higher degree of data security as default with encryption; autonomy to decide otherwise	saving data locally on a patient's device; absolute authority; encryption; complete anonymization
Non-Exclusion of Participants	adaptation to users' height	adaptation to users' height	adaptation to users' height and addressing additional senses (smell, hearing etc.) for visually/ physically impaired users
Danger of Discrimination	not included in design	not included in design, but with explanation	empathy training to reduce bias against smokers as additional training in

Table 8.8: Examples of Possible Design Solutions for Virtual Reality Applications (Paper 14).

To address the issue of excluding participants, the initial solution consisted of adapting the height of the head-mounted to the actual height of the user to avoid back strain while doing the task, and to allow participants doing the task while sitting in a (wheel)chair. Interestingly, the ethics workshops did not yet reveal an additional solution that can be implemented throughout this project from a practical point of view. However, for the regulative idea, it would be desirable to address all senses, including the smell of stimuli as well as more detailed haptic and auditory feedback to allow visually and physically impaired individuals to use the application. However, as implementing these design elements is quite complex and it is not clear whether a nonvisual version of the

AAT would have a beneficial effect for participants, we concluded that we could not implement them during the current project.

Concerning societal changes, neither the initial solution nor in the solution after the ethics workshop included design features to address them. However, in the ethics workshops, the issue of possible discrimination against smokers was identified. As this issue lay too far from the main goals of the project, it was concluded that an explanation about the issue and the missing inclusion of it in the VR-application should be included in the project reports related to the ethics workshops. Additionally, we concluded that some kind of empathy training to reduce intergroup bias against smokers for both users of the VR-application, as well as non-smokers, would constitute the regulative idea, which was out of the scope of the current project.

8.2.6 Results and Discussion

With our approach to document the design process and make ethical research more comparable, we contribute to addressing responsibility more transparently. Based on the identified table, the initial solution, the solution after the ethics workshop, and the "regulative idea" are comparable. On this basis, we can, at least qualitatively, conclude that the ethics workshops affected the refined solution. Additionally, we could make transparent that in all cases, a more beneficial regulative idea would have existed, which is not feasible to implement over the course of the current project. Our next steps lie in the development of a methodological approach with which the differences between the different solutions we identified can be analyzed not only qualitatively, but quantitatively with statistical methods and in comparison to other approaches, like the ValSD and VirtSD. For this, we need to clarify how we can operationalize a difference measure for the distance between the initial solution, the solution after ethics workshops, and the regulative idea and how these different measures could be applied in different contexts of responsible innovation research. Related to this, we have to decide if any additional ethical conflicts should be considered and which level of detail is reasonable to address them. In the current stage of our method, ethics workshops are conducted after the designers complete an initial design. It is important to note that with ethics workshops conducted before the first design, using a checklist, or designers were highly trained in ethics, differences between the first solution of the designers and the solution after additional ethics workshops might become smaller, which can still indicate effectiveness of the RRI approach. Therefore, future research can investigate whether there is a smaller difference in these cases compared with ethics workshops conducted at a later stage. All in all, we expect that our method contributes to enabling researchers to evaluate different RRI approaches quantitatively and subsequently, allows RRI researchers to identify the best approach for motivating designers to integrate ethical considerations into design in the long-term.

8.2.7 Acknowledgements

This paper is part of the research project ANTARES (promotional reference 005-1706-0006, Ministry for Culture and Science of the State North Rhine-Westphalia).

Part C: Synopsis

9 Results

9.1 Track 1: Gamification Design Elements

To gain insights into research question 1 ("How should gamification elements be designed to increase motivational and performance-related outcomes?") and research question 2 ("How can the effects of gamification design elements on motivational and performance-related outcomes be explained?"), this section summarizes the findings on the effects of gamification elements in traditional technology (paper 1, paper 3, and paper 4) and immersive technology (paper 2).

9.1.1 Traditional Technology

In paper 1, we conducted a 2 (avatar similarity: low versus high) × 2 (embodied feedback: no feedback versus embodied feedback) × 2 (status feedback: no feedback versus score and leaderboard) + 1 (control group) between-subjects experiment with a final sample of 332 participants recruited via a crowdworking platform. The results of the study show that gamification design elements can positively influence motivational outcomes. Figure 9.1 shows the structural equation model. The Fornell–Larcker criterion was mostly met (Table 9.1), and the model showed an adequate fit both with and without sociodemographic variables (Table 9.2).

Construct	1	2	3	4	5	6	7	8
1. Identification	**.87**							
2. Dia. Avatar	.70	**.85**						
3. Dia. Task	.39	.40	**.83**					
4. Relatedness	.66	.45	.40	**.86**				
5. Autonomy (Mean.)	.42	.41	.57	.51	**.80**			
6. Autonomy (Dec.)	.24	.11	.28	.25	.42	**.83**		
7. Competence	.40	.34	.47	.46	.79	.37	**.82**	
8. Reuse Intention	.41	.40	.44	.50	.81	.22	.54	**.84**

Table 9.1: Correlations of Constructs in the Solution of the Confirmatory Factor Analysis (Paper 1). The bold numbers display the square root of the average variance extracted.

The results of the structural equation model show that avatar similarity, status feedback, and embodied feedback influence identification, avatar diagnosticity, and task diagnosticity, which further influence reuse intention through the three psychological needs, identification, and diagnosticity perceptions. The need for autonomy can be further dif-

	Without Control Variables	With Control Variables
χ^2(df)	571.12 (312)	677.67(413)
CFI	.954	.953
RMSEA	.050	.044
SRMR	.055	.052
AIC	20034.28	23647.411

Table 9.2: Fit Indices of the Different Models (Paper 1)

ferentiated into autonomy for task meaningfulness and autonomy for decision freedom. Importantly, whereas autonomy for task meaningfulness has a positive effect on reuse intention, autonomy for decision freedom has a negative effect. Means and standard deviations are displayed in Table 9.3.

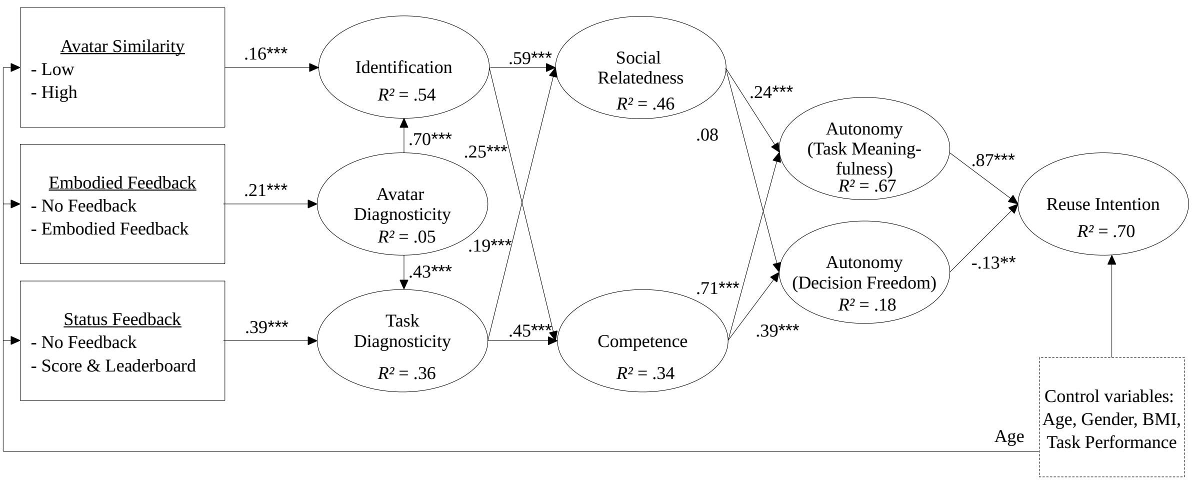


Figure 9.1: Results on the Effect of Embodied Feedback and Status Feedback on Reuse Intention (Paper 1)

Finally, the results highlight the important role that virtual bodies play in motivation and acceptance. Although paper 1 clearly shows that embodied feedback is a useful gamification design element and that designing individualized gamification elements with high similarity to the user is beneficial for motivation and performance, there are also hints toward a possible uncanny valley effect from the increased human likeness associated with high similarity. How these aspects may affect acceptance of virtual bodies is further investigated in track 2.

The studies in track 1 also show that gamification design elements can be implemented in IVR (paper 3), and current theoretical insights indicate a positive relationship to both motivational and performance-related aspects of behavior related to health.

		Low Similarity			
		Not Embodied		Embodied	
	Control	No Status	Status	No Status	Status
Identification	3.08 (1.46)	2.31 (1.43)	2.27 (1.30)	2.40 (1.43)	3.47 (1.74)
Dia. Avatar	4.09 (1.47)	2.93 (1.63)	2.89 (1.56)	3.96 (1.75)	4.37 (1.50)
Dia. Task	4.29 (1.58)	4.01 (1.60)	5.25 (1.32)	4.48 (1.41)	5.71 (1.04)
Relatedness	2.58 (1.57)	2.54 (1.43)	2.87 (1.56)	2.52 (1.38)	3.43 (1.72)
Competence	5.03 (1.25)	4.66 (1.12)	5.30 (1.39)	4.94 (1.22)	5.65 (0.93)
Autonomy (Mean.)	4.76 (1.40)	4.51 (1.43)	5.19 (1.28)	4.82 (1.32)	5.40 (1.15)
Autonomy (Deci.)	3.94 (1.97)	3.86 (1.75)	4.80 (1.80)	4.58 (1.73)	4.72 (1.59)
Reuse	4.46 (1.62)	4.09 (1.39)	4.44 (1.43)	4.16 (1.58)	4.94 (1.24)
		High Similarity			
		Not Embodied		Embodied	
		No Status	Status	No Status	Status
Identification		2.91 (1.46)	3.12 (1.57)	3.43 (1.36)	3.01 (1.58)
Dia. Avatar		3.36 (1.46)	2.83 (1.69)	4.24 (1.60)	3.65 (1.66)
Dia. Task		4.28 (1.65)	5.05 (1.57)	4.59 (1.20)	5.40 (1.23)
Relatedness		2.25 (1.35)	2.49 (1.38)	2.97 (1.50)	2.57 (1.50)
Competence		4.88 (1.39)	5.28 (1.05)	5.18 (1.07)	5.34 (1.03)
Autonomy (Mean.)		4.64 (1.20)	5.09 (1.37)	4.80 (1.43)	5.14 (1.20)
Autonomy (Deci.)		4.11 (2.13)	4.93 (1.65)	4.67 (1.53)	4.93 (1.80)
Reuse		4.28 (1.50)	4.55 (1.63)	4.14 (1.50)	4.34 (1.50)

Table 9.3: Means and Standard Deviations (in Brackets) for the Manipulation Checks and Dependent Variables (Paper 1)

Gamification can also be used outside the health domain, for example for facilitating knowledge sharing (paper 4). Here, design elements aimed at facilitating collaborations might be best suited for positive motivational and performance-related effects. Specifically, paper 4 shows that the effect of gamification design elements in the form of badges is likely to be positive in a non-health context. Furthermore, by using a work-related context highly embedded in pro-social behavior, it is shown that gamification design elements can also affect users' behavior through the knowledge that others will see their performance in a certain task.

9.1.2 Immersive Virtual Reality

This section summarizes the results of the investigation of gamification design elements in IVR by using proximal food cues as interaction objects in a gamified application in the context of health (paper 2). In this study, users' perception on whether objects were actually present in the virtual environment (object presence) was investigated. Specifically in relation to the approach avoidance task (AAT), the inclusion of proximal cues as 3D objects (high object presence) is more beneficial for these outcomes than using 2D objects (low object presence). The following describes the results of a 2 (object presence: low versus high) × 2 (training condition: avoiding versus approaching unhealthy food in the form of chocolate) experiment investigating the effects on enjoyment, craving for unhealthy behavior, and unhealthy behavior itself. The means and standard deviations for these results are shown in Table 9.4.

	High		Low	
Construct	Approach	Avoid	Approach	Avoid
1. Enjoyment	5.25 (0.54)	5.08 (0.57)	5.00 (0.35)	4.52 (0.81)
2. CEQ-S (Pre)	3.52 (1.07)	3.41 (1.12)	3.45 (1.14)	3.61 (1.39)
3. CEQ-S (Post)	3.39 (1.38)	3.7 (1.35)	3.04 (1.11)	3.07 (1.24)
4. Chocolate Consumption	15.03 (8.65)	16.01 (12.71)	18.33 (10.69)	20.32 (11.30)
5. Chocolate Bias (Pre)	-46.47 (216.57)	-83.03 (145.58)	-84.8 (114.22)	-63.55 (99.55)
6. Chocolate Bias (Post)	-9.76 (178.17)	-75.95 (104.19)	-67.42 (145.84)	-76.5 (73.63)
7. Fruit Bias (Pre)	-73.03 (143.97)	-76.65 (121.32)	-62.75 (104.57)	-123.73 (179.50)
8. Fruit Bias (Post)	-20.93 (112.69)	-40.42 (129.64)	-81.80 (125.24)	-18.80 (76.73)
9. Object Presence	5.54 (1.05)	5.55 (0.78)	5.02 (1.38)	4.5 (1.78)
10. Telepresence	5.41 (0.81)	4.6 (1.20)	5.12 (0.87)	4.71 (1.21)

Table 9.4: Means and Standard Deviations for Dependent Variables and Manipulation Checks (Paper 2)

Motivational outcomes. A 2 × 2 ANOVA on motivational outcomes in the form of enjoyment in doing the AAT in IVR showed that high object presence leads to higher enjoyment (mean (M) = 5.16, standard deviation (SD) = 0.56) than low object presence ($M = 4.76, SD = 0.67$, $p = .003$). Additionally, participants experienced higher enjoyment for approaching ($M = 5.12, SD = 0.47$) than for avoiding ($M = 4.79, SD = 0.75$) stimuli. The interaction effect was not significant ($F(2, 79) = 1.32, p = .252, \eta^2_G = 0.017$). Furthermore, a mixed 2 × 2 × 2 ANOVA revealed a significant time × object presence interaction effect ($F(1, 79) = 5.50, p = .022, \eta^2_G = 0.013$), shown in Figure 9.2. No other main or interaction effects were significant ($p > .102$ for all values). This interaction indicates that participants in the low object presence condition experienced lower craving after the training ($M = 3.06$, $SD = 1.16$) than before the training ($M = 3.53$, 1.26, $F(1,41) = 8.40, p = .006, \eta^2_G = 0.038$). The other main and interaction effects for low and high object presence did not reveal significant differences ($p > .225$ for all values). This suggests that training with high object presence can maintain high craving during and after training, whereas craving is reduced for low object presence.

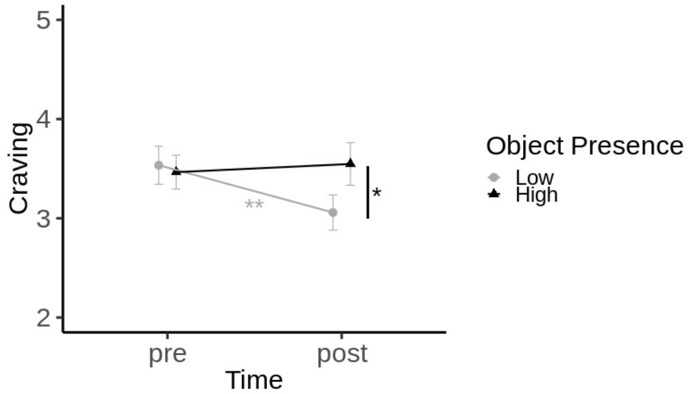


Figure 9.2: Interaction Effect for Craving (Paper 2)

Performance-related outcomes. For performance-related outcomes, an ANOVA for chocolate consumption found that participants in the low object presence condition ($M = 19.35g, SD = 10.93$) showed more unhealthy behavior in the form of eating chocolate than those in the high object presence condition ($M = 15.52g$, $SD = 10.74$, $F = 4.00, p = 0.049, eta^2_G = 0.048$). All other main and interaction effects revealed no statistical significance.

Approach bias. For chocolate bias, there was neither a main effect nor an interaction effect ($p > .297$ for all values). For fruit approach bias, only the main effect for time was significant ($F(1,78) = 5.80, p = .018, \eta^2_G = 0.027$). Thus, fruit approach bias scores were higher after training ($M = -39.96, SD = 113.01$) than before training ($M = -85.01, SD = 140.77$), regardless of the condition. This indicates that participants learned to show less avoidance for fruit stimuli after being exposed to the IVR training.

Paper 3 further investigated the role of gamification design elements in IVR for doing exercise by building a virtual fitness center. Here, the gamification design elements of leaderboards and avatars were expected to affect performance through the three psychological needs. Avatars, especially avatars placed in the same room as the user, instead of on a virtual television screen, were expected to enhance performance because of social facilitation effects due to increased co-presence.

9.2 Track 2: Virtual Bodies

To gain more insights into the design of virtual bodies, this section summarizes the findings for research question 3 ("What role does similarity play in motivational and performance-related outcomes?") and research question 4 ("What role do reflective and automatic processes in anthropomorphism play in the acceptance of virtual bodies?"). Overall, the results show support for a dual-process perspective of virtual bodies in terms of the visual similarity of virtual bodies to humans, which elicits anthropomorphism.

First, the results of a within-subjects experiment show that reuse intention of virtual bodies in the form of social robots is influenced by both reflective and impulsive trust. Furthermore, impulsive trust seems to influence use intention both directly as well as via mediation through reflective trust (Figure 9.3 and Table 9.5).

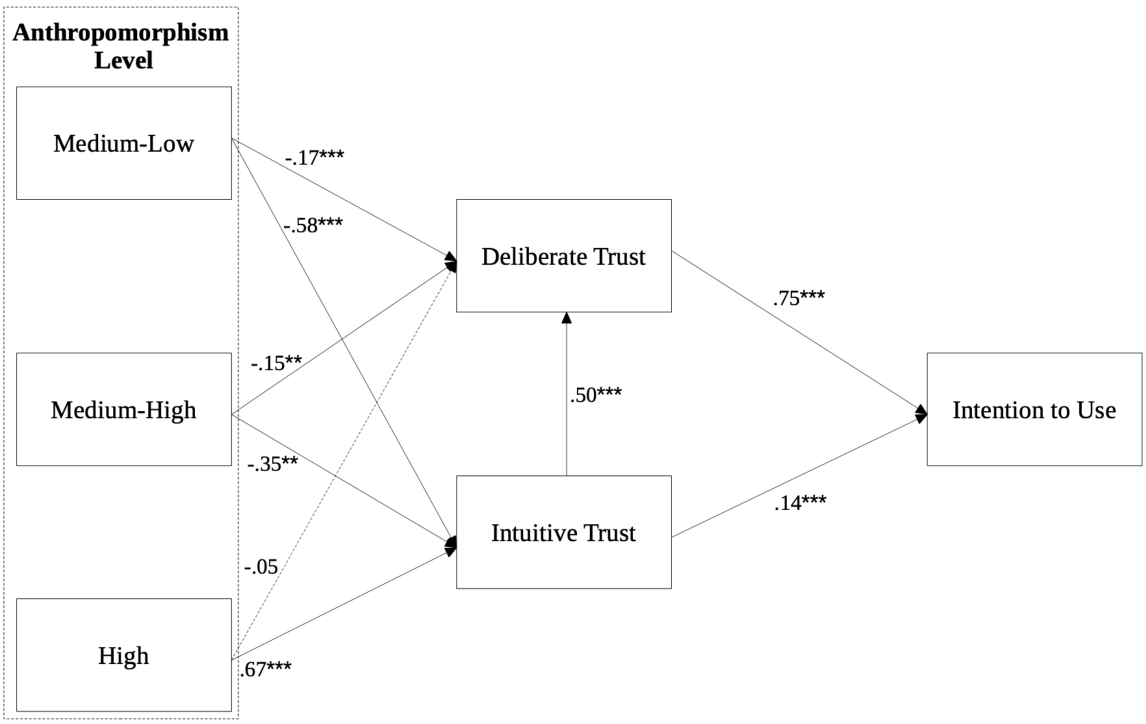


Figure 9.3: Results of the Mediation Model (Paper 5)

Second, the results of a 2 (robotic anthropomorphism: low versus high) × 2 (humanness: human versus robot) experiment show that there are distinctive effects for the impulsive and reflective system. On the one hand, the reflective system showed significantly reduced trust and warmth for robots compared with humans for the high and low robotic anthropomorphism condition. On the other hand, the impulsive system in the form of memory bias and categorization bias showed only significantly increased bias against

Anthropomorphism	Low M (SD)	Medium-Low M (SD)	Medium-High M (SD)	High/Human M (SD)
Intuitive PT	2.94 (1.08)	2.27 (0.91)	2.53 (1.03)	3.70 (1.03)
Deliberate PT	3.86 (1.51)	3.13 (1.47)	3.34 (1.51)	4.30 (1.53)
Use Intention	4.02 (1.60)	3.25 (1.59)	3.51 (1.54)	4.48 (1.56)

Table 9.5: Means and Standard Deviations (Paper 5)

robots for the low robotic anthropomorphism condition, but not for the high robotic anthropomorphism condition.

	Low Robotic Anthropomorphism		High Robotic Anthropomorphism	
	Robot	Human	Robot	Human
Memory Bias	7.42 (2.41)	6.28 (2.88)	7.35 (3.02)	7.26 (3.10)
Trust	4.12 (1.37)	4.64 (0.74)	3.89 (1.11)	4.80 (1.01)
Warmth	4.00 (1.31)	4.41 (0.61)	3.78 (0.99)	4.56 (0.86)

Table 9.6: Means and Standard Deviations (Paper 6)

Finally, the results of paper 7 show that IVR can likely be used to tackle the uncanny valley effect by embodying a user as a virtual body. Moreover, paper 8 also highlights cases in which using a virtual body dissimilar to the user's body can be more beneficial in terms of increasing users' performance.

9.3 Track 3: Gamification Contexts

This section looks at the different contexts in which gamification may take place. The findings from the research papers relating to work life and personal life are summarized below in answer to research question 5 ("How can learning in different contexts of individualized learning systems be explained?").

9.3.1 Work Life

In the area of work life, we selected the topic of maintaining work–life balance as a gamification context and conducted an interview study. We identified several technological implementations based on six individual tactics that could be gamified (Figure 9.4), which are described below.

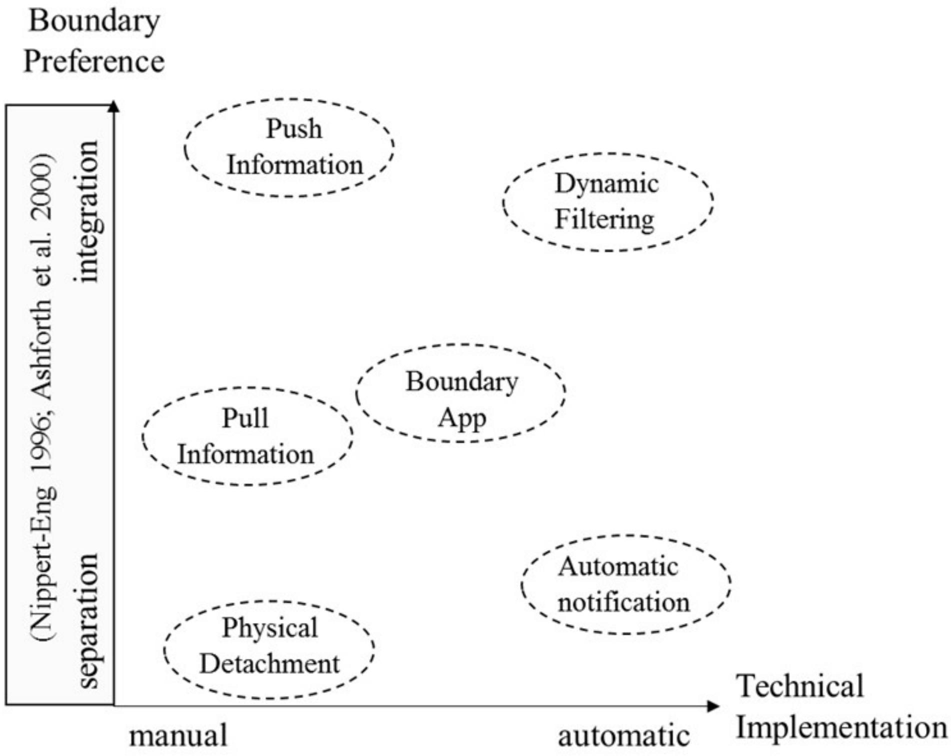


Figure 9.4: IT-related Boundary Tactics (Paper 10)

First, two individual tactics can facilitate separation of work and private life domains. On the one hand, the tactic "physical detachment" assists individuals by completely detaching them from all information, either by turning their smart devices off or leaving them at home or at the office – this tactic is by far the one with the lowest level of automaticity. On the other hand, by implementing automatic responses for messages, users can still use their devices, but will know that messages are taken care of by the system. Second, tactics that support integration use push information or dynamic filtering. Push information provides constant information about any incoming phone call or message. In contrast, dynamic filtering activates instant notifications only for specific messages or senders.

Additionally, we identified two tactics for mediation between integration and separation, namely pull information and a boundary app. Whereas pull information provides only a manual opportunity to look up messages or phone calls, thus providing a low level of automaticity, a boundary app could allow users to switch between work and personal life domains within the same technology. Therefore, by addressing individualized preferences through an adaptable interface, a boundary app provides a high level of individualization while maintaining a medium level of automaticity. How these different tactics can benefit from gamification toward aligning boundary preferences with actual boundary enactment is further elaborated on in the discussion section (sections 6.1 and 6.2).

9.3.2 Personal Life

The results of a survey with 320 participants on the topic of toxic behavior in online games show that toxic behavior is primarily influenced by toxic disinhibition. Additionally, attitudes toward toxic behavior and behavioral control influence toxic behavior directly, and show an indirect effect, together with toxic behavior victimization, through toxic disinhibition. How these results can be used to inform gamification research on reducing toxic behavior in online games is explained in the discussion section (section 6.1 and 6.2). The results are shown in Figure 9.5.

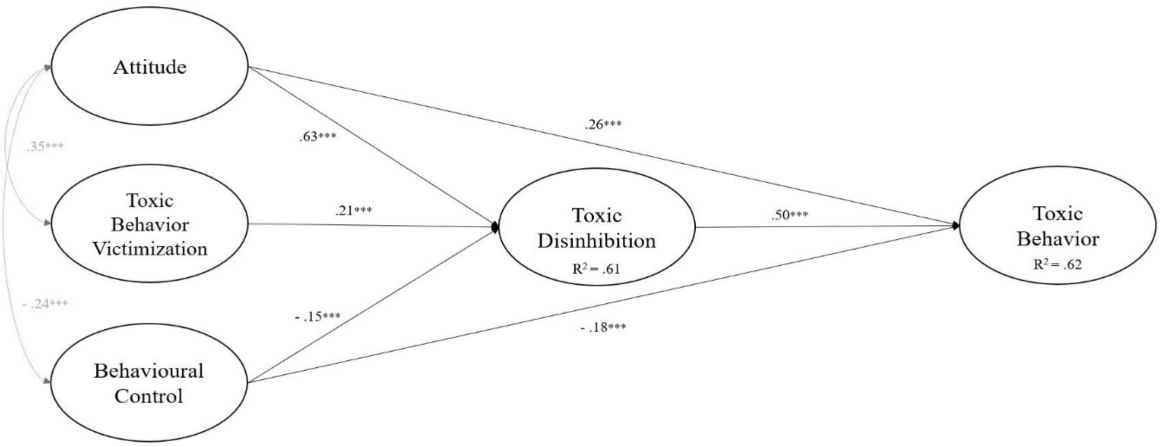


Figure 9.5: Final Path Model (Paper 9)

Paper 11 and paper 12 further show how learning can be facilitated through emerging technologies such as IVR and affective technologies. In IVR, by recreating the recall environment for the learning phase, environmental context-dependent effects can be used to enhance learning. Regarding gamification, a relevant question is how gamification design elements that are present in the learning but not in the recall phase might affect this process. Regarding affective technologies, transparency might play an important role in users' acceptance of the technology. Therefore, if gamification design elements are linked to the emotional state of the user, designers should be aware of the degree of transparency that the system delivers.

	determinants for acceptance	determinants for rejection
VR/game	future-orientation, usefulness, fun, curiosity, past experience, experiencing phantasies, physical activity, health	motion sickness, third party's unfamiliarity, distance from reality, mental-related changes, social-related changes, ugly design, lacking maturity, distraction from other activities, fear of addiction
VR/health care	physical enhancement, mental enhancement, usefulness, curiosity, future orientation, familiarity	lacking usefulness, lacking understanding of functionality, lacking trust in media/functionality/doctor, strain, technology hostility, motion sickness, addiction, high costs, time, fear, potential harm
affective technology	trust in functionality/ government/structural assurances, low costs, usefulness, just world beliefs, health, curiosity, social-related changes, physical enhancement, mental enhancement, technological advancement	lacking trust in functionality/-government/manufacturer, lacking usefulness, high costs, surveillance, "sick" behavior, fear of addiction, cheaper substitute, technology hostility

Table 9.7: Results of Steps 4 and 5 "Clustering and Identifying Determinants" (Paper 13).

9.4 Track 4: Ethical Implications

Finally, this section summarizes the findings regarding ethical implications of gamified IS by answering research question 6 ("How can ethical implications for unfamiliar technologies be investigated?") and research question 7 ("How can design choices with ethical implications be made transparent in the design process?"). To investigate ethical implications, we developed a method to identify morally relevant design implications that determine users' acceptance and rejection of emerging technologies, to facilitate upstream engagement, and applied these in a case study. In the RRI process, upstream engagement is used to gain initial information on the attitudes of potential stakeholders (Warburton et al., 2008). We developed the upstream engagement approach using text mining in the form of sentiment detection (Liu, 2010).

The developed method consists of six steps. First, suitable posts in social media platforms are selected. In the case study, we used Facebook as the social media platform and used the topics of affective technology and IVR as application cases. Second, com-

	determinant	design implication
VR/game	+experiencing fantasies	Use VR to bring fantasies alive (e.g. based on novels with a strong fan base)
	-social-related changes (isolation)	Meeting people in reality can be designed as a part of the game.
VR/health care	+physical enhancement	Use VR to teach doctors how to move correctly.
	-lacking understanding of/trust in functionality	The tool should provide sophisticated help and be designed "transparently" (what is it doing? which data does it collect? which dangers exist? etc.).
affective technology	+trust in structural assurances	Provide the users of affective technology with external safety structures (e.g. regulations, laws, guarantees etc.).
	-surveillance	Do not forward information about the user or, if necessary, make transparent who gets which information.

Table 9.8: Examples of Derived Design Implications for IVR and Affective Technology (Paper 13)

ments are categorized into acceptance, rejection, and neutral opinions. In the third step, the reasons for acceptance and rejection are extracted, and in the fourth step, they are clustered into relevant superordinate categories. In the fifth step, the determinants for acceptance and rejection are identified by linking the clusters to relevant theories. Finally, in the sixth step, design implications are derived.

The results of the application of this method in the context of two emerging technologies, namely IVR and affective technology, are shown in Table 9.7 and Table 9.8. For IVR in the gaming context, a determinant for acceptance is the opportunity to experience fantasies. From this, it can be derived that one possible way to facilitate acceptance of IVR games is to bring fantasies alive (e.g., in the form of novels that can be experienced in IVR). In contrast, a determinant for rejection consists of expecting negative social-related changes (e.g., isolation). As a result, designing IVR with an opportunity to meet users in the real world could be a meaningful design implication to address this aspect. Examples of determinants for IVR in health care are opportunities for physical enhancement (acceptance) or reduced understanding and trust (rejection). Finally, for affective technology, determinants are trust in structural assurances (acceptance) and surveillance (rejection).

In paper 14, we also developed a method to increase transparency of design decisions in RRI, using a case study, and to facilitate evaluation of different RRI approaches. In this method, specific norms are identified that are relevant for the design of the product from an ethical viewpoint. Subsequently, by using ethics workshops, the initial design solution, which is provided by the developers, the solution proposed following ethics workshops, and a regulative idea are recorded. Whereas the initial solution and the solution proposed following ethics workshops consist of designs that can actually be applied, the regulative idea is the most desirable design that may or may not be applied in the actual design of the application (e.g., because of limited resources). The solutions can then be qualitatively compared, which allows measurement of the effectiveness of the applied RRI approach. An example of the development of an IVR application for health purposes is given in Table 9.9.

Norm	Initial Solution	Solution after the Ethics Workshops	Regulative Idea
Privacy	saving data pseudonymized on VR-computer	higher degree of data security as default with encryption; autonomy to decide otherwise	saving data locally on a patient's device; absolute authority; encryption; complete anonymization
Non-Exclusion of Participants	adaptation to users' height	adaptation to users' height	adaptation to users' height and addressing additional senses (smell, hearing etc.) for visually/ physically impaired users
Danger of Discrimination	not included in design	not included in design, but with explanation	empathy training to reduce bias against smokers as additional training

Table 9.9: Examples of Possible Design Solutions for Virtual Reality Applications (Paper 14)

10 Discussion

10.1 Contribution to Theory

The findings presented in this thesis contribute to the development of a gamification-related learning theory. Specific contributions to the literature in terms of the research questions are as follows.

*Ad RQ*1. How should gamification elements be designed to increase motivational and performance-related outcomes?

Research track 1 found several implications for the design of gamification elements. First, the finding that increased avatar similarity is beneficial for motivational outcomes supports the personalization principle (Liu et al., 2017). This is in line with previous research that found a favorable effect of avatars created through body scanning (Suh et al., 2011). Further, we found that even a relatively vague degree of user similarity, achieved by matching reported gender, can be enough to increase positive motivational outcomes. Future research in IS can build on this result to investigate the role of similarity in a diverse set of gamification contexts.

Second, the finding that specific embodied feedback designed to relate closely to the task facilitates motivational outcomes supports the task congruence principle (Liu et al., 2017). Future research can build on this result by investigating the degree to which it is important to match the implementation of the feedback to the consequences of the behavior that should be learned. For example, an application aimed at teaching good work–life balance tactics may require a different type of embodied feedback than an application aimed at reducing toxic behavior. Here, designing experiments comparing different kinds of embodied feedback within and across different contexts could provide more theoretical and practical insights into this aspect.

Third, the finding that object presence increases both motivational and performance-related outcomes provides initial insights into the design of gamified IVR applications. Here, specifically in relation to the transfer of cognitive tasks traditionally used in psychological interventions, such as the AAT (Machulska et al., 2016; Schumacher et al., 2016; Wiers et al., 2010), it seems to be more beneficial to design stimuli with a high degree of object presence rather than designing stimuli with lesser object presence but higher similarity to the original AAT conducted at a desktop computer. Future research in IS can benefit from investigating the topic of object presence in relation to other technologies, for example with respect to mixed or augmented technology.

*Ad RQ*2. How can the effects of gamification design elements on motivational and performance-related outcomes be explained?

Track 1 found that gamification effects can be explained by self-determination theory (Deci & Ryan, 1985) and approaches to presence (Schultze, 2010). For self-determination theory, gamification design elements influence motivational outcomes through the three psychological needs of competence, social relatedness, and autonomy. Moreover, differentiating autonomy into autonomy for task meaningfulness and autonomy for decision freedom provides unique explanatory power. Regarding presence, the findings show that both object presence and co-presence may affect motivational and performance-related outcomes and that these effects can be explained by self-determination theory.

*Ad RQ*3. What role does similarity play in motivational and performance-related outcomes?

Track 2 was aimed at identifying the role of user similarity in motivational and performance-related outcomes and was investigated by increasing the human likeness of virtual bodies in paper 6, paper 5, and paper 7. The findings show that increased human likeness can indeed lead to the uncanny valley effect (Mori, 1970). Therefore, the adoption of virtual robots in the form of avatars or social robots could be a hindrance because predictors and consequences of motivational outcomes, such as trust and acceptance, are reduced. Furthermore, paper 8 found that scenarios exist in which reduced similarity to users can in fact be beneficial for performance-related outcomes.

*Ad RQ*4. What role do reflective and automatic processes in anthropomorphism play in the acceptance of virtual bodies?

Paper 5 found that uncanny valley effects may persist both in the automatic and reflective systems. In addition, the automatic system also affects the reflective system (see 10.1). As a result, design elements aimed at mitigating these effects should be aimed at both the automatic and reflective system. However paper 6 found that not all aspects of the automatic system are affected by the uncanny valley effect. The findings of paper 7 show that there are opportunities in IVR to investigate the effects of human likeness of virtual bodies in the automatic system in contexts that have greater ecological validity. Finally, the results of paper 8 highlight the relevance of investigating the effects of (dis)similarity to the user in relation to NeuroIS measures.

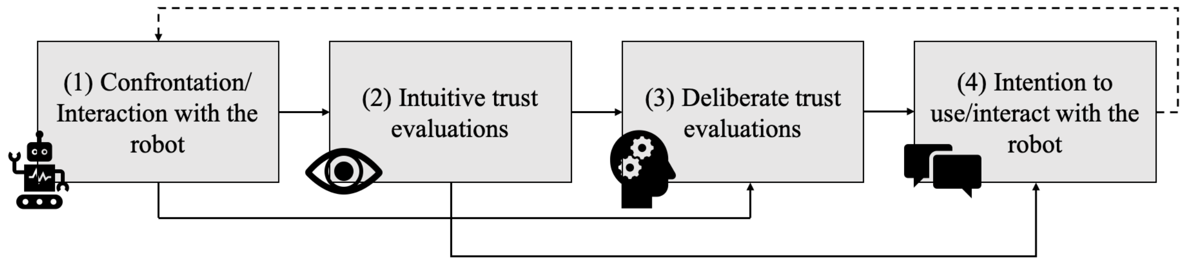


Figure 10.1: A Dual-Processing Model of Perceived Trust (Paper 5)

*Ad RQ*5. How can learning in different contexts of individualized learning systems be explained?

Looking at the implications of different contexts of individualized learning systems, research track 3 found that the contexts in which gamification could be implemented to facilitate learning are diverse. By using the unified theory of toxic behavior, future gamification research could design specific gamification elements that help to reduce toxic behavior. Furthermore, the identification of IT-related boundary tactics informs gamification research on behaviors that could be facilitated by using certain gamification design elements. In relation to users' boundary preferences (Kreiner, 2006), gamification design elements that facilitate tactics that contribute to strengthening the enactment of boundary management according to these preferences can be developed and evaluated.

*Ad RQ*6. How can ethical implications for unfamiliar technologies be investigated?

In track 4, research question 4 was aimed at investigating the ethical implications of technologies with which stakeholders may be unfamiliar. The proposed method contributes to facilitating upstream engagement by reducing the effort needed to engage participants, resulting in temporal and monetary advantages over traditional approaches. Therefore, particularly research at an early stage of product development can benefit from this approach. In the gamification domain, this method could be used to investigate new gamification applications, particularly for emerging technologies such as IVR, AR, or affective technology.

*Ad RQ*7. How can design choices with ethical implications be made transparent in the design process?

Track 4 also aimed to investigate the transparency of design choices with ethical implications in the design process. By using the proposed method, research on responsible innovation, particularly in the form of value-sensitive design (Friedman & Hendry, 2019) and virtue-sensitive design (Vallor, 2016) can benefit from being able to assess the effectiveness of different responsible innovation approaches as well as different implementations of responsible innovation approaches. Furthermore, by making transparent what design variants best address a specific norm in the form of regulative ideas, designers in the RRI area can, in the long term, work more effectively toward such alternatives. As a consequence, the development of gamification applications could benefit from applying this approach from the beginning of the product development until the final design, with the goals of progressively identifying RRI approaches that a) identify the best regulative ideas and b) are most effective at revealing the design variants that most closely match the regulative ideas. As an example, the method could be used to further validate the effectiveness of the method developed in paper 13 regarding more economic upstream engagement.

10.2 Contribution to Real-World Practice

This thesis contributes to real-world practice in several ways. First, designers can benefit from the insights offered into the selection of gamification design elements for individualized learning systems. In particular, individualized learning systems should be designed with a high degree of individualization. Second, in relation to individualization, the findings show that for virtual bodies used in gamified learning systems, there might be a trade-off between the degree of individualization in the form of human likeness of virtual bodies and the uncanny valley effect. Here, designers should make sure that they either use an almost perfect human-like design or – if this is not feasible – use a design that has a relatively low degree of human likeness. Third, the thesis shows that designers should carefully consider the design of gamification elements in the context of the learning system, particularly in relation to embodied feedback. While in the context of health applications, one way seems to be to use virtual bodies that adapt to the choices the users make in a health-relevant aspect (e.g., gaining/losing weight in the context of eating food), it is likely that other contexts could also benefit from a matching design. For example, in the context of reducing toxic behavior, it might be beneficial to use embodied feedback with a smiling versus angry avatar according to the level of toxic behavior. In contrast, in the context of teaching good work–life balance, it would be necessary to assess users' preferences regarding their desired level of work–life balance and subsequently use an avatar with a stressed or relaxed face according to their success in achieving the desired work–life balance through different boundary tactics. However, clearly, more research is needed to assess how important contextual matching is across different contexts, given that it might be that any level of embodied feedback, regardless of contextual fit, is beneficial. Fourth, individualized learning systems in IVR should be designed with a high level of object presence to maximize beneficial learning effects in both reflective and impulsive systems. Finally, designers should take into account the ethical implications of their systems by applying an RRI approach with a high level of transparency regarding their design decisions to facilitate long-term evaluation of RRI methods, which may help to increase acceptance and acceptability of their design artifact.

10.3 Limitations and Future Research

This thesis is not without limitations. However, given that the specific limitations are elaborated on in the individual papers, this section focuses on the limitations in a broader sense. First, the papers investigated only a selection of contexts and applications. Regarding application contexts, the papers focused in particular on health applications and the adoption of robots without any interaction. Consequently, it is still unclear to which degree similar effects would arise in other health domains as well as in robot adoption scenarios after an initial interaction.

Second, all studies used a convenient sample, which consisted mostly of students or crowdworkers in the German context. Although the studies investigated effects that are highly likely to be consistent over different groups, it might be that differences between specific groups could be revealed by future research, for example with regard to cultural effects. Additionally, the sample size of some studies could have been higher. This might have led to some studies being unable to detect all possible effects. Therefore, future studies should investigate these research questions by using a higher sample size or within-subjects designs to increase their power.

Third, studies that investigated mediation effects used designs in which only the independent variable was varied experimentally, not the mediators. To rule out other possible explanations of how the mediation effect might have arisen, future studies could investigate more complex mediation designs, for example by trying to block the mediator (MacKinnon & Fairchild, 2009). Moreover, future studies could test mediation effects with additional statistical approaches such as sensitivity analyses (Imai et al., 2010; Imai & Yamamoto, 2013).

Fourth, most of the studies used a quantitative approach, and no mixed method studies were applied. Mixed method studies have a unique ability to provide insights into emerging topics about which little is known. Thus, investigating gamified IS from a mixed methods perspective could bring additional insights into the working mechanisms of gamification design elements, particularly with relatively new technology becoming available as tool for gamified IS, in the form of increasingly elaborate digital agents.

Fifth, only a few of the papers applied a NeuroIS approach, in which processes in the brain can be investigated in a more objective way. Exceptions include paper 2, which used the AAT and paper 6, which applied the "who said what" paradigm, both of which can reveal associations in the brain of participants. Consequently, investigating uncanny valley effects by using biophysiological measures, such as skin conductance response, and neuropsychological measures, such as functional near-infrared spectroscopy, as suggested in paper 8 and paper 6, would be a sensible route for future research.

Finally, all quantitative papers used a cross-sectional approach, investigating effects only once or occurring within one hour. Therefore, it remains an open question whether and how design elements change motivational and performance-related outcomes over a time span of weeks or months. Moreover, it is unclear whether attitudes in the context of work–life balance, social robots, and toxic behavior in video games remain stable over time. These topics could be investigated in future research.

11 Conclusion

This thesis investigated the effects of gamification elements on users' motivational and performance-related outcomes, the contexts in which gamification can be applied, and relevant ethical implications for the design of gamification applications. The results of research track 1 show that using individualized gamification design elements can enhance motivational and performance-related outcomes. However, the results of track 2 also show that when individualization is designed for virtual bodies, it also comes with the pitfall of reducing motivational outcomes due to the uncanny valley effect because increased similarity to the user can also come with increased human likeness. Furthermore, this effect may differ in the automatic and reflective system. As a result, future research should keep in mind possible positive and negative effects that may arise due to specific gamification design elements and relate their results to dual-process theory. Moreover, in investigating different gamification contexts in track 3, this thesis shows that learning may take place differently in different contexts, learning goals can be diverse, and opportunities for learning specific contexts should be considered in the design of suitable gamification design elements. Finally, the investigation of ethical implications in track 4 reveals specific suggestions for how to implement RRI as early as possible in the design process, how to evaluate its effectiveness continuously, and how to report results in an informative way that facilitates future research, by explicitly referring to regulative ideas. As a result, future research can build on these findings to investigate the effects of gamification design elements regarding different contexts, long-term effects, and NeuroIS measurements. This thesis also calls for ethical implications to be considered as early as possible in the design and with increased transparency, to evaluate the effectiveness of the different implementations of RRI.

References

Adinolf, S., & Turkay, S. (2018). Toxic behaviors in esports games: Player perceptions and coping strategies. *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, 365–372.

Afzal, S., & Robinson, P. (2015). Emotion data collection and its implication for affective computing. In R. A. Calvo, S. K. D'Mello, J. Gratch, & A. Kappas (Eds.), *The oxford handbook of affective computing* (pp. 359–370). Oxford University Press.

Agarwal, R., Gao, G. G., DesRoches, C., & Jha, A. K. (2010). The digital transformation of healthcare: Current status and the road ahead. *Information Systems Research*, *21*(4), 796–809.

Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly*, *24*(4), 665–694.

Ahrens, W., Pigeot, I., Pohlabeln, H., De Henauw, S., Lissner, L., Molnár, D., Moreno, L. A., Tornaritis, M., Veidebaum, T., & Siani, A. (2014). Prevalence of overweight and obesity in european children below the age of 10. *International Journal of Obesity*, *38*, 99–107.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Prentice-Hall.

Alavi, M. (1994). Computer-mediated collaborative learning: An empirical evaluation. *MIS Quarterly*, *18*(2), 159–174.

Alavi, M., & Leidner, D. E. (2001). Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, *25*(1), 107–136.

Albrechtslund, A. (2007). Ethics and technology design. *Ethics and Information Technology*, *9*(1), 63–72.

Allen, T. D., Cho, E., & Meier, L. L. (2014). Work–family boundary dynamics. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*(1), 99–121.

Altschuller, S., & Benbunan-Fich, R. (2013). The pursuit of trust in ad hoc virtual teams: How much electronic portrayal is too much? *European Journal of Information Systems*, *22*(6), 619–636.

Ammons, S. K. (2013). Work-family boundary strategies: Stability and alignment between preferred and enacted boundaries. *Journal of Vocational Behavior*, *82*(1), 49–58.

Arai, S., Sakamoto, K., Washizaki, H., & Fukazawa, Y. (2014). A gamified tool for motivating developers to remove warnings of bug pattern tools. *2014 6th International Workshop on Empirical Software Engineering in Practice*, 37–42.

Arnold, V., Clark, N., Collier, P. A., Leech, S. A., & Sutton, S. G. (2006). The differential use and effect of knowledge-based system explanations in novice and expert judgment decisions. *MIS Quarterly*, *30*(1), 79–97.

Ashforth, B. E., Kreiner, G. E., & Fugate, M. (2000). All in a day's work: Boundaries and micro role transitions. *The Academy of Management Review*, *25*(3), 472–491.

Bailenson, J. N. et al. (2005). The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, *14*(4), 379–393.

Baker, W. E., & Bulkley, N. (2014). Paying it forward vs. rewarding reputation: Mechanisms of generalized reciprocity. *Organization Science*, *25*(5), 1493–1510.

Baldwin, N., Branyon, J., Sethumadhavan, A., & Pak, R. (2015). In search of virtual social facilitation effects. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *59*(1), 90–94.

Banakou, D., Hanumanthu, P. D., & Slater, M. (2016). Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias. *Frontiers in Human Neuroscience*, *10*, 601.

Banakou, D., Kishore, S., & Slater, M. (2018). Virtually being Einstein results in an improvement in cognitive task performance and a decrease in age bias. *Frontiers in Psychology*, *9*, 917.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ, US: Prentice-Hall, Inc.

Bandura, A. (2002). Social cognitive theory in cultural context. *Applied Psychology*, *51*(2), 269–290.

Baptista, G., & Oliveira, T. (2017). Why so serious? Gamification impact in the acceptance of mobile banking services. *Internet Research*, *27*(1), 118–139.

Barlett, C. P., Gentile, D. A., & Chew, C. (2016). Predicting cyberbullying from anonymity. *Psychology of Popular Media Culture*, *5*(2), 171–180.

Bar-On, R., & Parker, J. D. A. (2000). *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace*. Jossey-Bass.

Bartneck, C., Yogeeswaran, K., Ser, Q. M., Woodward, G., Sparrow, R., Wang, S., & Eyssel, F. (2018). Robots and racism. *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 196–204.

Bartol, K. M., & Srivastava, A. (2002). Encouraging knowledge sharing: The role of organizational reward systems. *Journal of Leadership & Organizational Studies*, *9*(1), 64–76.

Baskerville, R., & Pries-Heje, J. (2010). Explanatory design theory. *Business & Information Systems Engineering*, *2*(5), 271–282.

Bastiaensens, S., Pabian, S., Vandebosch, H., Poels, K., Van Cleemput, K., DeSmet, A., & De Bourdeaudhuij, I. (2016). From normative influence to social pressure: How relevant others affect whether bystanders join in cyberbullying: Normative influence on cyberbullying bystanders. *Social Development*, *25*(1), 193–211.

Beard, C., Weisberg, R. B., & Primack, J. (2012). Socially anxious primary care patients' attitudes toward cognitive bias modification (CBM): A qualitative study. *Behavioural and Cognitive Psychotherapy*, *40*(5), 618–633.

Becker, D., Jostmann, N. B., Wiers, R. W., & Holland, R. W. (2015). Approach avoidance training in the eating domain: Testing the effectiveness across three single session studies. *Appetite*, *85*, 58–65.

Beiboer, J., & Sandoval, E. B. (2019). Validating the accuracy of imaged-based research into the uncanny valley: An experimental proposal. *ACM/IEEE International Conference on Human-Robot Interaction*, *2019-March*, 608–609.

Bekele, E., Bian, D., Peterman, J., Park, S., & Sarkar, N. (2017). Design of a virtual reality system for affect analysis in facial expressions (VR-SAAFE); application to schizophrenia. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *25*(6), 739–749.

Belpaeme, T., Vogt, P., van den Berghe, R., Bergmann, K., Göksun, T., de Haas, M., Kanero, J., Kennedy, J., Küntay, A. C., Oudgenoeg-Paz, O., Papadopoulos, F., Schodde, T., Verhagen, J., Wallbridge, C. D., Willemsen, B., de Wit, J., Geçkin, V., Hoffmann, L., Kopp, S., … Pandey, A. K. (2018). Guidelines for designing social robots as second language tutors. *International Journal of Social Robotics*, *10*(3), 325–341.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate : A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Bergmann, K., Eyssel, F., & Kopp, S. (2012). A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7502*, 126–138.

Berking, M., & Znoj, H. (2008). Entwicklung und validierung eines fragebogens zur standardisierten selbsteinschätzung emotionaler kompetenzen (SEK-27). *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, *56*(2), 141–153.

Bessant, J. (2013). Innovation in the twenty-first century. *Responsible innovation - managing the responsible emergence of science and innovation in society* (p. 14). John Wiley & Sons Ltd.

Betsch, C., & Kunz, J. J. (2008). Individual strategy preferences and decisional fit. *Journal of Behavioral Decision Making*, *21*(5), 532–555.

Birk, M. V., Atkins, C., Bowey, J. T., & Mandryk, R. L. (2016). Fostering intrinsic motivation through avatar identification in digital games. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, 2982–2995.

Blackburn, J., & Kwak, H. (2014). STFU NOOB! Predicting crowdsourced decisions on toxic behavior in online games. *Proceedings of the 23rd International Conference on World Wide Web*, 877–888.

Blanz, M. (2015). *Forschungsmethoden der Statistik für die Soziale Arbeit: Grundlagen und Anwendungen*. Kohlhammer.

Bock, Zmud, Kim, & Lee. (2005). Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. *MIS Quarterly*, *29*(1), 87–111.

Bockelman Morrow, P., & Fiore, S. M. (2012). Supporting human-robot teams in social dynamicism: An overview of the metaphoric inference framework. *Proceedings of the Human Factors and Ergonomics Society*, 1718–1722.

Bojd, B., Song, X., Tan, Y., & Yan, X. (2018). Gamified goals: An empirical study of online weight-loss challenges. *Proceedings of the 39th International Conference on Information Systems (ICIS), San Francisco, USA*.

Bordnick, P. S., Carter, B. L., & Traylor, A. C. (2011). What virtual reality research in addictions can tell us about the future of obesity assessment and treatment. *Journal of Diabetes Science and Technology*, *5*(2), 265–271.

Bostrom, R. P., Olfman, L., & Sein, M. K. (1990). The importance of learning style in end-user training. *MIS Quarterly*, *14*(1), 101–119.

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature*, *391*, 756.

Boyatzis, R., Goleman, D., & Rhee, K. (2000). Clustering competence in emotional intelligence: Insights from the emotional competence inventory (ECI). In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace* (pp. 343–362). Jossey-Bass Publishers.

Broadbent, E., Lee, Y. I., Stafford, R. Q., Kuo, I. H., & MacDonald, B. A. (2011). Mental schemas of robots as more human-like are associated with higher blood pressure and negative emotions in a human-robot interaction. *International Journal of Social Robotics*, *3*(3), 291–297.

Bujić, M., Salminen, M., Macey, J., & Hamari, J. (2020). "Empathy machine": How virtual reality affects human rights attitudes. *Internet Research*, *30*(5), 1407–1425.

Burton, P., Smit, H. J., & Lightowler, H. J. (2007). The influence of restrained and external eating patterns on overeating. *Appetite*, *49*(1), 191–197.

Burton-Jones, A., & Straub, D. W. (2006). Reconceptualizing system usage: An approach and empirical test. *Information Systems Research*, *17*(3), 228–246.

Butler, A. C., Godbole, N., & Marsh, E. J. (2013). Explanation feedback is better than correct answer feedback for promoting transfer of learning. *Journal of Educational Psychology*, *105*(2), 290–298.

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2007). *Handbook of psychophysiology*. Cambridge University Press.

Canas, J. J., & Nelson, D. L. (1986). Recognition and environmental context: The effect of testing by phone. *Bulletin of the Psychonomic Society*, *24*(6), 407–409.

Carvalhaes, M. F. A., Neto, O. C. d. S., Ferreira, J. O., Rocha, A. F. d., & Barbosa, T. M. G. d. A. (2013). Including affectivity requirements in embedded systems. *2013 IEEE International Systems Conference (SysCon)*, 229–234.

Cecchinato, M. E., Cox, A. L., & Bird, J. (2015). Working 9-5?: Professional differences in email and boundary management practices. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3989–3998.

Cerasoli, C. P., Nicklin, J. M., & Nassrelgrgawi, A. S. (2016). Performance, incentives, and needs for autonomy, competence, and relatedness: A meta-analysis. *Motivation and Emotion*, *40*(6), 781–813.

Chaminade, T., Zecca, M., Blakemore, S. J., Takanishi, A., Frith, C. D., Micera, S., Dario, P., Rizzolatti, G., Gallese, V., & Umiltà, M. A. (2010). Brain response to a humanoid robot in areas implicated in the perception of human emotional gestures. *PLoS ONE*, *5*(7), e11577.

Chatfield, T. (2011). *Fun inc.: Why gaming will dominate the twenty-first century*. Pegasus Books.

Chatman, J. A. (1989). Improving interactional organizational research: A model of person-organization fit. *The Academy of Management Review*, *14* (3), 333.

Chen, C., Zhang, K. Z., Gong, X., & Lee, M. (2019). Dual mechanisms of reinforcement reward and habit in driving smartphone addiction: The role of smartphone features. *Internet Research*, *29* (6), 1551–1570.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, *25* (2), 215–224.

Chu, C.-W., & Lu, H.-P. (2007). Factors influencing online music purchase intention in taiwan: An empirical study based on the value-intention framework. *Internet Research*, *17* (2), 139–155.

Ciarrochi, J. V., & Deane, F. P. (2001). Emotional competence and willingness to seek help from professional and nonprofessional sources. *British Journal of Guidance & Counselling*, *29* (2), 233–246.

Clark, S. C. (2000). Work/family border theory: A new theory of work/family balance. *Human Relations*, *53* (6), 747–770.

Codish, D., & Ravid, G. (2017). Gender moderation in gamification: Does one size fit all? *Proceedings of the 50th Hawaii International Conference on System Sciences, HICSS-50, Maui, Hawaii, USA*, 2006–2015.

Constant, D., Kiesler, S., & Sproull, L. (1994). What's mine is ours, or is it? A study of attitudes about information sharing. *Information Systems Research*, *5* (4), 400–421.

Coppola, N. W., Hiltz, S. R., & Rotter, N. G. (2002). Becoming a virtual professor: Pedagogical roles and asynchronous learning networks. *Journal of Management Information Systems*, *18* (4), 169–189.

Corbin, J. M., & Strauss, A. L. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Thousand Oaks.

Cousijn, J., Goudriaan, A. E., & Wiers, R. W. (2011). Reaching out towards cannabis: Approach-bias in heavy cannabis users predicts changes in cannabis use. *Addiction (Abingdon, England)*, *106* (9), 1667–1674.

Cousins, K., & Robey, D. (2015). Managing work-life boundaries with mobile technologies - an interpretive study of mobile work practices. *Information Technology & People*, *28*, 34–71.

Cowie, R. (2015). Ethical issues in affective computing. In R. A. Calvo, S. K. D'Mello, J. Gratch, & A. Kappas (Eds.), *The oxford handbook of affective computing* (p. 334.). Oxford University Press.

Csikszentmihalyi, M. (1990). *Flow - the psychology of optimal experience*. Harper Perennial Modern Classics.

Culley, K. E., & Madhavan, P. (2013). A note of caution regarding anthropomorphism in HCI agents. *Computers in Human Behavior*, *29*(3), 577–579.

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, *19*(2), 272–309.

Dadgar, M., & Joshi, K. D. (2018). The role of information and communication technology in self-management of chronic diseases: An empirical investigation through value sensitive design. *Journal of the Association for Information Systems*, *19*(2), 86–112.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of somputer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982–1003.

Davis, G. B. (2002). Anytime/anyplace computing and the future of knowledge work. *Communications of the ACM*, *45*(12), 67–73.

Davis, J., & Nathan, L. P. (2015). Value sensitive design: Applications, adaptations, and critiques. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.), *Handbook of ethics, values, and technological design: Sources, theory, values and application domains* (pp. 11–40). Springer Netherlands.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *Catalog of Selected Documents in Psychology*, *10*.

De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic Bulletin & Review*, *20*(4), 631–642.

De Houwer, J., Crombez, G., Baeyens, F., & Hermans, D. (2001). On the generality of the affective Simon effect. *Cognition & Emotion*, *15*(2), 189–206.

de Brujin, G., Kremers, S., Singh, A., Van den Putte, B., & Van Mechelen, W. (2009). Adult active transportation: Adding habit strength to the theory of planned behavior. *American Journal of Preventive Medicine*, *36*, 189–194.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer Science+Business Media.

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*, *11*(4), 227–268.

de Mesquita Neto, J. A., & Becker, K. (2018). Relating conversational topics and toxic behavior effects in a MOBA game. *Entertainment Computing*, *26*, 10–29.

den Hamer, A. H., & Konijn, E. A. (2015). Adolescents' media exposure may increase their cyberbullying behavior: A longitudinal study. *Journal of Adolescent Health*, *56*(2), 203–208.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification. using game-design elements in non-gaming contexts. *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '11*.

Deutsch, R., & Strack, F. (2006). Duality models in social psychology: From dual processes to interacting systems. *Psychological Inquiry*, *17*(3), 166–172.

DeVellis, R. (2016). *Scale development: Theory and applications* (26th ed.). Sage Publications.

Dibbets, P., & Fonteyne, R. (2015). High spider fearfuls can overcome their fear in a virtual approach-avoidance conflict task. *Journal of Depression & Anxiety*, *4*(2), 1–7.

Dickson, H., Kavanagh, D. J., & MacLeod, C. (2016). The pulling power of chocolate: Effects of approach-avoidance training on approach bias and consumption. *Appetite*, *99*, 46–51.

Dimoka, A., Banker, R. D., Benbasat, I., Davis, F. D., Dennis, A. R., Gefen, D., Gupta, A., Ischebeck, A., Kenning, P., Müller-Putz, G., Pavlou, P. A., Riedl, R., vom Brocke, J., & Weber, B. (2012). On the use of neurophysiological tools in IS research: Developing a research agenda for NeuroIS. *MIS Quarterly*, *36*(3), 679–702.

Dimoka, A., Pavlou, P. A., & Davis, F. D. (2010). Research commentary—NeuroIS: The potential of cognitive neuroscience for information systems research. *Information Systems Research*, *22*(4), 687–702.

Dingle, G. A., Dingle, G. A., Gleadhill, L., & Baker, F. A. (2008). Can music therapy engage patients in group cognitive behaviour therapy for substance abuse treatment? *Drug and Alcohol Review*, *27*(2), 190–196.

DiSalvo, C. F., Gemperle, F., Forlizzi, J., & Kiesler, S. (2002). All robots are not created equal: The design and perception of humanoid robot heads. *Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, DIS*, 321–326.

DiStefano, C., Zhu, M., & Mîndril, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research and Evaluation*, *14*(20), 1–11.

D'Mello, S., & Calvo, R. A. (2013). Beyond the basic emotions: What should affective computing compute? *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, 2287–2294.

D'Mello, S., & Graesser, A. (2013). AutoTutor and affective autotutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems*, *2*(4), 23:1–23:39.

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153–170.

Doane, A. N., Pearson, M. R., & Kelley, M. L. (2014). Predictors of cyberbullying perpetration among college students: An application of the theory of reasoned action. *Computers in Human Behavior*, *36*, 154–162.

Doney, P. M., Cannon, J. P., & Mullen, M. R. (1998). Understanding the influence of national culture on the development of trust. *Academy of Management Review*, *23*(3), 601–620.

Dovidio, J. F., Gaertner, S. L., Validzic, A., Matoka, K., Johnson, B., & Frazier, S. (1997). Extending the benefits of recategorization: Evaluations, self-disclosure, and helping. *Journal of Experimental Social Psychology*, *33*(4), 401–420.

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, *42*(3), 177–190.

Dulewicz, V., & Higgs, M. (2000). Emotional intelligence – a review and evaluation study. *Journal of Managerial Psychology*, *15*(4), 341–372.

Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2013). *Fundamentals of business process management*. Springer.

Duxbury, L., Higgins, C., Smart, R., & Stevenson, M. (2014). Mobile technology and boundary permeability. *British Journal of Management*, 19.

Ellemers, N., Kortekaas, P., & Ouwerkerk, J. W. (1999). Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. *European Journal of Social Psychology*, *29*(23), 371–389.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115.

Elson, J., Derrick, D. C., & Ligon, G. S. (2020). Trusting a humanoid robot: Exploring personality and trusting effects in a human-robot partnership. *Proceedings of the 53rd Hawaii International Conference on System Sciences, Maui, Hawaii, USA*, 543–553.

Emmerich, K., & Masuch, M. (2016). The influence of virtual agents on player experience and performance. *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886.

Ershow, A. G., Peterson, C. M., Riley, W. T., Rizzo, A. S., & Wansink, B. (2011). Virtual reality technologies for research and education in obesity and diabetes: Research needs and opportunities. *Journal of Diabetes Science and Technology*, *5*(2), 212–224.

Ewoldsen, D. R., Eno, C. A., Okdie, B. M., Velez, J. A., Guadagno, R. E., & DeCoster, J. (2012). Effect of playing violent video games cooperatively or competitively on subsequent cooperative behavior. *Cyberpsychology, Behavior, and Social Networking*, *15*(5), 277–280.

Eyssel, F., & Kuchenbrandt, D. (2012). Social categorization of social robots: Anthropomorphism as a function of robot group membership. *British Journal of Social Psychology*, *51*(4), 724–731.

Faiola, A., Newlon, C., Pfaff, M., & Smyslova, O. (2013). Correlating the effects of flow and telepresence in virtual worlds: Enhancing our understanding of user behavior in game-based learning. *Computers in Human Behavior*, *29*(3), 1113–1121.

Fair Play Alliance. (2018). Retrieved April 26, 2018, from http://www.fairplayalliance.org

Falconer, C. J., Rovira, A., King, J. A., Gilbert, P., Antley, A., Fearon, P., Ralph, N., Slater, M., & Brewin, C. R. (2016). Embodying self-compassion within virtual reality and its effects on patients with depression. *British Journal of Psychiatry Open*, *2*(1), 74–80.

Fan, H., & Lederman, R. (2018). Online health communities: How do community members build the trust required to adopt information and form close relationships? *European Journal of Information Systems*, *27*(1), 62–89.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175–191.

Fehrenbacher, D. D. (2017). Affect infusion and detection through faces in computer-mediated knowledge-sharing decisions. *Journal of the Association for Information Systems*, *18*(10), 703–726.

Fernandes, J., Duarte, D., Ribeiro, C., Farinha, C., Pereira, J. M., & Silva, M. M. d. (2012). iThink: A game-based approach towards improving collaboration and participation in requirement elicitation. *Procedia Computer Science*, *15*, 66–77.

Fernández, Á., & Alonso, M. A. (2001). The relative value of environmental context reinstatement in free recall. *Psicológica*, *22*(2), 253–266.

Ferrer-Garcia, M., Gutierrez-Maldonado, J., Treasure, J., & Vilalta-Abella, F. (2015). Craving for food in virtual reality scenarios in non-clinical sample: Analysis of its relationship with body mass index and eating disorder symptoms. *European Eating Disorders Review: The Journal of the Eating Disorders Association*, *23*(5), 371–378.

Ferrer-Garcia, M., Pla-Sanjuanelo, J., Dakanalis, A., Vilalta-Abella, F., Riva, G., Fernandez-Aranda, F., Forcano, L., Riesco, N., Sánchez, I., Clerici, M., Ribas-Sabaté, J., Andreu-Gracia, A., Escandón-Nagel, N., Gomez-Tricio, O., Tena, V., & Gutiérrez-Maldonado, J. (2019). A randomized trial of virtual reality-based cue exposure second-level therapy and cognitive behavior second-level therapy for bulimia nervosa and binge-eating disorder: Outcome at six-month followup. *Cyberpsychology, Behavior and Social Networking*, *22*(1), 60–68.

Ferrer-García, M., Gutiérrez-Maldonado, J., & Riva, G. (2013). Virtual reality based treatments in eating disorders and obesity: A review. *Journal of Contemporary Psychotherapy*, *43*, 207–221.

Ffiske, T. (2020). *Educational VR: The best learning apps to try out (2020)*. Retrieved December 11, 2020, from https://www.virtualperceptions.com/educational-vr/%20(accessed%2015%20November%202020)

Fisher, E., Mahajan, R. L., & Mitcham, C. (2006). Midstream modulation of technology: Governance from within. *Bulletin of Science, Technology & Society*, *26*(6), 485–496.

Fleck, R., Cox, A. L., & Robison, R. A. (2015). Balancing boundaries: Using multiple devices to manage work-life balance. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3985–3988.

Foerster, J., Grant, H., Idson, L. C., & Higgins, E. (2001). Success/failure feedback, expectancies, and approach/avoidance motivation: How regulatory focus moderates classic relations. *Journal of Experimental Social Psychology*, *37*(3), 253–260.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolotti, G. (2005). Neuroscience: Parietal lobe: From action organization to intention understanding. *Science*, *308*(5722), 662–667.

Fornell, C., & Larcker, D. (1981a). *Structural equation models with unobservable variables and measurement error: Algebra and statistics*. SAGE Publications.

Fornell, C., & Larcker, D. F. (1981b). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.

Fox, J., Bailenson, J., & Binney, J. (2009). Virtual experiences, physical behaviors: The effect of presence on imitation of an eating avatar. *Presence: Teleoperators and Virtual Environments*, *18*(4), 294–303.

Fox, J., & Bailenson, J. N. (2009). Virtual self-modeling: The effects of vicarious reinforcement and identification on exercise behaviors. *Media Psychology*, *12*(1), 1–25.

Fox, J., & Tang, W. Y. (2014). Sexism in online video games: The role of conformity to masculine norms and social dominance orientation. *Computers in Human Behavior*, *33*, 314–320.

French, J. R., Caplan, R. D., & Van Harrison, R. (1982). *The mechanisms of job stress and strain*. John Wiley & Sons Ltd.

Friedman, B., & Hendry, D. G. (2019). *Value sensitive design: Shaping technology with moral imagination*. The MIT Press.

Friedman, B., Hendry, D. G., & Borning, A. (2017). A survey of value sensitive design methods. *Foundations and Trends® in Human–Computer Interaction*, *11*(2), 63–125.

Friedman, B., Kahn, P. H., & Borning, A. (2008). Value sensitive design and information systems. *The handbook of information and computer ethics*, 69–101.

Fukuyama, F. (1996). *Trust: the social virtues and the creation of prosperity* (1. Free Press paperback ed). Free Press.

Funk, D. C., Pizzo, A. D., & Baker, B. J. (2018). eSport management: Embracing eSport education and research opportunities. *Sport Management Review*, *21*(1), 7–13.

Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (1993). The common ingroup identity model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology*, *4*(1), 1–26.

Gaertner, S. L., Mann, J., Murrell, A., & Dovidio, J. F. (1989). Reducing intergroup bias: The benefits of recategorization. *Journal of Personality and Social Psychology*, *57*(2), 239–249.

García-Rodríguez, O., Weidberg, S., Gutiérrez-Maldonado, J., & Secades-Villa, R. (2013). Smoking a virtual cigarette increases craving among smokers. *Addictive Behaviors*, *38*(10), 2551–2554.

Gefen, D. (2000). E-commerce: The role of familiarity and trust. *Omega*, *28*(6), 725–737.

Gefen, D., Karahanna, E., & Straub, D. (2003). Trust and TAM in online shopping: An integrated model. *MIS Quarterly*, *27*(1), 51–90.

Gefen, D., Straub, D., & Boudreau, M.-C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems*, *4*, 7.

Gethmann, C. F., & Sander, T. (1999). Rechtfertigungsdiskurse. In A. Grunwald & S. Saupe (Eds.), *Ethik in der Technikgestaltung: Praktische Relevanz und Legitimation* (pp. 117–151). Springer Berlin Heidelberg.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research* (4. paperback printing). Aldine.

Glöckner, A., & Witteman, C. (2009). *Foundations for tracing intuition: Challenges and methods*. Taylor & Francis.

Gockley, R., Bruce, A., Forlizzi, J., Michalowski, M., Mundell, A., Rosenthal, S., Sellner, B., Simmons, R., Snipes, K., Schultz, A. C., & Wang, J. (2005). Designing robots for long-term social interaction. *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, 2199–2204.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of psychology*, *66*(3), 325–331.

Goldman, A. (2008). Company on the couch: Unveiling toxic behavior in dysfunctional organizations. *Journal of Management Inquiry*, *17*(3), 226–238.

Goleman, D. (1999). *Working with emotional intelligence*. Ted Smart.

Gong, L. (2008). The boundary of racial prejudice: Comparing preferences for computer-synthesized white, black, and robot characters. *Computers in Human Behavior*, *24*(5), 2074–2093.

Gorini, A., Capideville, C. S., De Leo, G., Mantovani, F., & Riva, G. (2011). The role of immersion and narrative in mediated presence: The virtual hospital experience. *Cyberpsychology, Behavior and Social Networking*, *14*(3), 99–105.

Gorini, A., Griez, E., Petrova, A., & Riva, G. (2010). Assessment of the emotional responses produced by exposure to real food, virtual food and photographs of food in patients affected by eating disorders. *Annals of General Psychiatry*, *9*, 30.

Graesser, A., Chipman, P., Haynes, B., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, *48*(4), 612–618.

Green, J. E. (2014). Toxic leadership in educational organizations. *Education Leadership Review*, *15*(1), 18–33.

Gregor, S. (2009). Building theory in the sciences of the artificial. *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*, 1–10.

Groening, C., & Binnewies, C. (2019). Achievement unlocked! The impact of digital achievements as a gamification element on motivation and performance. *Computers in Human Behavior*, *97*, 151–166.

Grunwald, A. (1999). Technology assessment or ethics of technology? Reflections on technology development between social sciences and philosophy. *Ethical Perspectives*, *6*, 170–182.

Guo, Z., Xiao, L., Van Toorn, C., Lai, Y., & Seo, C. (2016). Promoting online learners' continuance intention: An integrated flow framework. *Information & Management*, *53*(2), 279–295.

Gutiérrez-Maldonado, J., Pla-Sanjuanelo, J., & Ferrer-García, M. (2016). Cue-exposure software for the treatment of bulimia nervosa and binge eating disorder. *Psicothema*, (28), 363–369.

Hall, D. T., & Richter, J. (1988). Balancing work life and home life: What can organizations do to help? *Academy of Management Perspectives*, *2*(3), 213–223.

Hamari, J. (2013). Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service. *Electronic Commerce Research and Applications*, *12*(4), 236–245.

Hamari, J., & Sjöblom, M. (2017). What is eSports and why do people watch it? *Internet Research*, *27*(2), 211–232.

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., de Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *53*(5), 517–527.

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, *80*, 152–161.

Harman, H. (1960). *Modern factor analysis*. University of Chicago Press.

Hartmann, E. (1991). *Boundaries in the mind*. Basic Books.

Hasler, B. S., Spanlang, B., & Slater, M. (2017). Virtual race transformation reverses racial ingroup bias. *Plos One*, *12*(4), 1–20.

Hayashi, K., Shiomi, M., Kanda, T., & Hagitay, N. (2010). Who is appropriate? a robot, human and mascot perform three troublesome tasks. *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, 348–354.

Heger, O., Jahn, K., Mueller, M., & Niehaves, B. (2017). Making use of facebook comments for upstream engagement: A systematic approach. *CEPE/ETHICOMP 2017. Turin, Italy.*

Heger, O., Kampling, H., & Niehaves, B. (2016). Towards a theory of trust-based acceptance of affective technology. *Twenty-Fourth European Conference on Information Systems (ECIS), İstanbul,Turkey.*

Hehman, E., Ingbretsen, Z. A., & Freeman, J. B. (2014). The neural basis of stereotypic impact on multiple social categorization. *NeuroImage*, *101*, 704–711.

Heirman, W., & Walrave, M. (2012). Predicting adolescent perpetration in cyberbullying: An application of the theory of planned behavior. *Psicothema*, *24*, 614–620.

Hew, K. F., Huang, B., Chu, K. W. S., & Chiu, D. K. W. (2016). Engaging asian students through game mechanics: Findings from two experiment studies. *Computers & Education*, *92-93*, 221–236.

Ho, S. S., Chen, L., & Ng, A. P. Y. (2017). Comparing cyberbullying perpetration on social media between primary and secondary school students. *Computers & Education*, *109*, 74–84.

Hoffman, H. G., Prothero, J., Wells, M. J., & Groen, J. (1998). Virtual chess: Meaning enhances users' sense of presence in virtual environments. *International Journal of Human–Computer Interaction*, *10*(3), 251–263.

Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science*, *4*(2), 162–176.

Hofmann, W., Friese, M., & Wiers, R. W. (2008). Impulsive versus reflective influences on health behavior: A theoretical framework and empirical review. *Health Psychology Review*, *2*(2), 111–137.

Hone-Blanchet, A., Wensing, T., & Fecteau, S. (2014). The use of virtual reality in craving assessment and cue-exposure therapy in substance use disorders. *Frontiers in Human Neuroscience*, *8*.

Hooff, B. v. d., & Ridder, J. A. d. (2004). Knowledge sharing in context: The influence of organizational commitment, communication climate and cmc use on knowledge sharing. *Journal of Knowledge Management*, *8*(6), 1367–3270.

Hosmer, L. T. (1995). Trust: The connecting link between organizational theory and philosophical ethics. *Academy of Management Review*, *20*(2), 379–403.

Hoyt, C. L., Blascovich, J., & Swinth, K. R. (2003). Social Inhibition in Immersive Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *12*(2), 183–195.

Hsu, C.-L., & Lu, H.-P. (2007). Consumer behavior in online game communities: A motivational factor perspective. *Computers in Human Behavior*, *23*(3), 1642–1659.

Hsu, M.-H., & Chiu, C.-M. (2004). Predicting electronic service continuance with a decomposed theory of planned behaviour. *Behaviour & Information Technology*, *23*(5), 359–373.

Hsu, M.-H., Ju, T. L., Yen, C.-H., & Chang, C.-M. (2007). Knowledge sharing behavior in virtual communities: The relationship between trust, self-efficacy, and outcome expectations. *International Journal of Human-Computer Studies*, *65*(2), 153–169.

Hsu, Y.-C., Tsai, C.-H., Kuo, Y.-M., & Ya-Hui, B. (2016). Telecare services for elderly: Predictive factors of continued use intention. *The Open Biomedical Engineering Journal*, *10*(1), 82–90.

Hupont, I., Baldassarri, S., Cerezo, E., & Del-Hoyo, R. (2013). The emotracker: Visualizing contents, gaze and emotions at a glance. *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 751–756.

Hur, I., Cousins, K. C., & Stahl, B. C. (2019). A critical perspective of engagement in online health communities. *European Journal of Information Systems*, *28*(5), 523–548.

IDC. (2015). *International data cooperation: Forecasts U.S. mobile worker population to surpass 105 million by 2020*. Retrieved April 8, 2016, from https://www.idc.com/getdoc.jsp?containerId=prUS25705415

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–334.

Imai, K., & Yamamoto, T. (2013). Identification and sensitivity analysis for multiple causal mechanisms: Revisiting evidence from framing experiments. *Political Analysis*, *21*(2), 141–171.

Inocencio-Gray, J., & Mercado, B. (2013). Toward a model of cyberbullying in the workplace: An online disinhibition perspective. *Academy of Management Annual Meeting Proceedings*, *2013*, 12352.

Isarida, T., & Isarin, T. K. (2007). Environmental context effects of background color in free recall. *Memory & Cognition*, *35*(7), 1620–1629.

Isarida, T., & Isarida, T. K. (2014). Environmental context-dependent memory. *Advances in experimental psychology research*, 115–151.

Isarida, T., Sakai, T., Kubota, T., Koga, M., Katayama, Y., & Isarida, T. K. (2014). Odor-context effects in free recall after a short retention interval: A new methodology for controlling adaptation. *Memory & Cognition*, *42*(3), 421–433.

Isarida, T. K., Kubota, T., Nakajima, S., & Isarida, T. (2017). Reexamination of mood-mediation hypothesis of background-music-dependent effects in free recall. *Quarterly Journal of Experimental Psychology (2006)*, *70*(3), 533–543.

Iversen, O. S., Halskov, K., & Leong, T. W. (2010). Rekindling values in participatory design. *Proceedings of the 11th Biennial Participatory Design Conference*, 91–100.

Jackson, L. A., von Eye, A., Biocca, F. A., Barbatsis, G., Zhao, Y., & Fitzgerald, H. E. (2006). Does home internet use influence the academic performance of low-income children? *Developmental Psychology*, *42*(3), 429–435.

Jackson, R., Barbagallo, F., & Haste, H. (2005). Strengths of public dialogue on science-related issues. *Critical Review of International Social and Political Philosophy*, *8*(3), 349–358.

Jacobs, N., & Huldtgren, A. (2018). Why value sensitive design needs ethical commitments. *Ethics and Information Technology*, *23*, 23–26.

Jahn, K. (2019). Gamification elements in immersive virtual reality. Comparing the effectiveness of leaderboards and copresence for motivation. *Proceedings of the 9th International Conference on Advanced Collaborative Networks, Systems and Applications (COLLA),* 8–11*, Rome, Italy.*

Jahn, K., Heger, O., Kampling, H., Stanik, K., & Niehaves, B. (2017). Designing for knowledge-based familiarity, trust, and acceptance: The case of affective technology. *Proceedings of the 25th European Conference on Information Systems (ECIS 2017), Guimarães, Portugal.*

Jahn, K., Kampling, H., Klein, H.-C., Kuru, Y., & Niehaves, B. (2018). Towards an explanatory design theory for context-dependent learning in immersive virtual reality. *Proceedings of the 22th Pacific Asia Conference on Information Systems (PACIS 2018), Yokohama, Japan.*

Jahn, K., Kempt, H., Eiler, T. J., Heger, O., Gruenewald, A., Mach, A., Klucken, T., Gethmann, C. F., Brueck, R., & Niehaves, B. (2020). More than ticking off a checklist? Towards an approach for quantifying the effectiveness of responsible innovation in the design process. *Workshop Ethik und Moral in der Wirtschaftsinformatik 2020, Potsdam, Germany.*

Jahn, K., Klesel, M., Lemmer, K., Weigel, A., & Niehaves, B. (2016). Individual boundary management: An empirical investigation on technology-related tactics. *Proceedings of the 20th Pacific Asia Conference on Information Systems (PACIS 2016), Chiayi, Taiwan.*

Jahn, K., & Kordyaka, B. (2019). The effects of robotic embodiment on intergroup bias: An experiment in immersive virtual reality. *Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm, Sweden.*

Jahn, K., Kordyaka, B., Machulska, A., Eiler, T. J., Gruenewald, A., Klucken, T., Brueck, R., Gethmann, C. F., & Niehaves, B. (2021). Individualized gamification elements: The impact of avatar and feedback design on reuse intention. *Computers in Human Behavior*, *119*, 106702.

Jahn, K., Kordyaka, B., Ressing, C., Roeding, K., & Niehaves, B. (2019). Designing self-presence in immersive virtual reality to improve cognitive performance - A research proposal. *Proceedings of the NeuroIS Retreat 2019, Vienna, Austria.*

Jahn, K., Kordyaka, B., Scholz, T., & Niehaves, B. (2020). Gamified helping? The impact of individualized and group-level cooperative evaluation on knowledge sharing. *Proceedings of the 15th Wirtschaftsinformatik 2020 (WI2020), Potsdam, Germany.*

Jahn, K., & Nissen, A. (2020). Towards dual processing of social robots: Differences in the automatic and reflective system. *Proceedings of the 41st International Conference on Information Systems (Virtual ICIS), Hyderabad, India.*

Jahn, K., Oschinsky, F., Kordyaka, B., Machulska, A., Eiler, T. J., Grünewald, A., Brück, R., Klucken, T., Gethmann, C. F., & Niehaves, B. (in revision). Design elements in immersive virtual reality: The impact of object presence on health-related outcomes. *Internet Research.*

Jessup, S. A., Gibson, A. M., Capiola, A., Alarcon, G. M., & Borders, M. (2020). Investigating the effect of trust manipulations on affect over time in human-human versus human-robot interactions. *Proceedings of the 53rd Hawaii International Conference on System Sciences, Maui, Hawaii, USA*, 553–563.

Jiang, Z., & Benbasat, I. (2007). Investigating the influence of the functional mechanisms of online product presentations. *Information Systems Research*, *18*(4), 454–470.

Jo, D., Kim, K., Welch, G. F., Jeon, W., Kim, Y., Kim, K.-H., & Kim, G. J. (2017). The impact of avatar-owner visual similarity on body ownership in immersive virtual reality. *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology - VRST '17*, 1–2.

Johnsen, K., & Lok, B. (2008). An evaluation of immersive displays for virtual human experiences. *2008 IEEE Virtual Reality Conference*, 133–136.

Johnson, D., Nacke, L. E., & Wyeth, P. (2015). All about that base: Differing player experiences in video game genres and the unique case of MOBA games. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2265–2274.

Kamlah, W., & Lorenzen, P. (1967). *Logische Propädeutik.* Bibliografisches Institut.

Kampling, H. (2018). The role of immersive virtual reality in individual learning. *Proceedings of the 51st Hawaii International Conference on System Sciences*, 1397–1406.

Kang, N., Ding, D., Riemsdijk, M. B. V., Morina, N., Neerincx, M. A., & Brinkman, W.-P. (2021). Self-identification with a virtual experience and its moderating effect on self-efficacy and presence. *International Journal of Human–Computer Interaction*, *37*(2), 181–196.

Kankanhalli, Tan, & Wei. (2005). Contributing knowledge to electronic knowledge repositories: An empirical investigation. *MIS Quarterly*, *29*(1), 113.

Karanasios, S., & Allen, D. (2014). Mobile technology in mobile work: Contradictions and congruencies in activity systems. *European Journal of Information Systems*, *23*(5), 529–542.

Karimi, J., Gupta, Y. P., & Somers, T. M. (1996). Impact of competitive strategy and information technology maturity on firms' strategic response to globalization. *Journal of Management Information Systems*, *12*(4), 55–88.

Katmada, A., Chatzakis, M., Apostolidis, H., Mavridis, A., & Panagiotis, S. (2015). An adaptive serious neuro-game using a mobile version of a bio-feedback device. *2015 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL)*, 416–420.

Kätsyri, J., Förger, K., Mäkäräinen, M., & Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology*, *6*, 390.

Katz, D., & Kahn, R. L. (1978). *The social psychology of organizations* (2nd ed). John Wiley & Sons Inc.

Kerr, I. R., & Bornfreund, M. (2005). Buddy bots: How Turing's fast friends are undermining consumer privacy. *Presence: Teleoperators and Virtual Environments*, *14*(6), 647–655.

Kilteni, K., Bergstrom, I., & Slater, M. (2013). Drumming in immersive virtual reality: The body shapes the way we play. *IEEE transactions on visualization and computer graphics*, *19*(4), 597–605.

Kilteni, K., Maselli, A., Kording, K. P., & Slater, M. (2015). Over my fake body: Body ownership illusions for studying the multisensory basis of own-body perception. *Frontiers in Human Neuroscience*, *9*, 141.

Kim, D.-Y., & Lee, J.-H. (2015). Development of a virtual approach-avoidance task to assess alcohol cravings. *Cyberpsychology, Behavior and Social Networking*, *18*(12), 763–766.

Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., & Scassellati, B. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of Autism and Developmental Disorders*, *43*(5), 1038–1049.

Kim, S., Na, E.-K., & Ryu, M.-H. (2007). Factors affecting user participation in video UCC (user-created contents) services. In C. Steinfield, B. T. Pentland, M. Ackerman, & N. Contractor (Eds.), *Communities and technologies 2007* (pp. 209–224). Springer London.

Kim, Y. J., & Shute, V. J. (2015). The interplay of game elements with psychometric qualities, learning, and enjoyment in game-based assessment. *Computers & Education*, *87*, 340–356.

King, N. (1998). *Qualitative methods in organizational research: A practical guide*. Template Analysis.

Kirby, R., Forlizzi, J., & Simmons, R. (2010). Affective social robots. *Robotics and Autonomous Systems*, *58*(3), 322–332.

Kishore, S., Muncunill, X. N., Bourdin, P., Or-Berkers, K., Friedman, D., Slater, M., & Slater, M. (2016). Multi-destination beaming: Apparently being in three places at once through robotic and virtual embodiment. *Frontiers in Robotics and AI*, *3*, 65.

Kline, R. (2015). *Principles and practice of structural equation modeling*. Guilford Publications.

Koeffer, S., Anlauf, L., Ortbach, K., & Niehaves, B. (2015). The intensified blurring of boundaries between work and private life through IT consumerization. *Proceedings of the 23rd European Conference on Information Systems, Münster, Germany.*

Kokkinara, E., & Slater, M. (2014). Measuring the effects through time of the influence of visuomotor and visuotactile synchronous stimulation on a virtual body ownership illusion. *Perception*, *43*(1), 43–58.

Komiak, S. X., & Benbasat, I. (2003). Understanding customer trust in agent-mediated electronic commerce, web-mediated electronic commerce, and traditional commerce. *Information Technology and Management*, *5*(1), 181–207.

Komiak, S. Y. X., & Benbasat, I. (2006). The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly*, *30*(4), 941–960.

Kordyaka, B., & Hribersek, S. (2019). Crafting identity in league of legends – Purchases as a tool to achieve desired impressions. *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS-52), Maui, Hawaii, USA*, 1506–1515.

Kordyaka, B., Jahn, K., & Niehaves, B. (2020). Towards a unified theory of toxic behavior in video games. *Internet Research*, *30*(4), 1081–1102.

Kordyaka, B., Klesel, M., & Jahn, K. (2019). Perpetrators in League of Legends: Scale development and validation of toxic behavior. *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS-52), Maui, Hawaii, USA*.

Kossek, E. E., & Lautsch, B. A. (2008). Work–family boundary management styles in organizations: A cross-level model. *Organizational Psychology Review*, *2*(2), 152–171.

Kossek, E. E., Ruderman, M. N., Braddy, P. W., & Hannum, K. M. (2012). Work–nonwork boundary management profiles: A person-centered approach. *Journal of Vocational Behavior*, *81*(1), 112–128.

Kouloumpis, E., Wilson, T., & Moore, J. (2011). Twitter sentiment analysis: The good the bad and the OMG! *Proceedings of the International AAAI Conference on Web and Social Media*, *5*(1), 538–541.

Kouper, I. (2010). Science blogs and public engagement with science: Practices, challenges, and opportunities. *Journal of Science Communication*, *09*(1), A02.

Kowalski, R., Giumetti, G., Schroeder, A., & Lattanner, M. (2014). Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth. *Psychological Bulletin*, *140*, 1073–1137.

Kozinets, R. V. (2002). The field behind the screen: Using netnography for marketing research in online communities. *Journal of Marketing Research*, *39*(1), 61–72.

Kozinets, R. V. (2010). *Netnography: Doing ethnographic research online* (1st ed.). Sage Publications.

Krach, S., Hegel, F., Wrede, B., Sagerer, G., Binkofski, F., & Kircher, T. (2008). Can machines think? Interaction and perspective taking with robots investigated via fMRI. *PLoS ONE*, *3*(7), e2597.

Kreiner, G. E. (2006). Consequences of work-home segmentation or integration: A person-environment fit perspective. *Journal of Organizational Behavior*, *27*(4), 485–507.

Kreiner, G. E., Hollensbe, E. C., & Sheep, M. L. (2009). Balancing borders and bridges: Negotiating the work-home interface via boundary work tactics. *Academy of Management Journal*, *52*(4), 704–730.

Krieglmeyer, R., & Deutsch, R. (2010). Comparing measures of approach–avoidance behaviour: The manikin task vs. two versions of the joystick task. *Cognition and Emotion*, *24*(5), 810–828.

Krikorian, R., Bartik, J., & Gay, N. (1994). Tower of london procedure: A standard method and developmental data. *Journal of Clinical and Experimental Neuropsychology*, *16*(6), 840–850.

Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology*, *49*(1), 1–49.

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

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, *1*(4), 279–308.

Kumar, N. (1996). The power of trust in manufacturer-retailer relationships. *Harvard Business Review*, *74*(6), 92–106.

Kumar, N., Scheer, L. K., & Steenkamp, J.-B. E. M. (1995). The effects of perceived interdependence on dealer attitudes. *Journal of Marketing Research*, *32*(3), 348–356.

Kummer, N., Kadish, D., Dulic, A., & Najjaran, H. (2012). The empathy machine. *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2265–2271.

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, *98*(26), 15387–15392.

Kwak, H., Blackburn, J., & Han, S. (2015). Exploring cyberbullying and other toxic behavior in team competition online games. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3739–3748.

Lafay, L., Thomas, F., Mennen, L., Charles, M. A., Eschwege, E., Borys, J. M., & Basdevant, A. (2001). Gender differences in the relation between food cravings and mood in an adult community: Results from the fleurbaix laventie ville santé study. *The International Journal of Eating Disorders*, *29*(2), 195–204.

Landers, R. N., & Landers, A. K. (2015). An empirical test of the theory of gamified learning: The effect of leaderboards on time-on-task and academic performance. *Simulation & Gaming*, *45*(6), 769–785.

Landowska, A. (2013). Affect-awareness framework for intelligent tutoring systems. *2013 6th International Conference on Human System Interactions (HSI)*, 540–547.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the radboud faces database. *Cognition & Emotion*, *24*(8), 1377–1388.

Lankton, N. K., Harrison McKnight, D., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. *Journal of the Association for Information Systems*, *16*(10), 880–918.

Lapidot-Lefler, N., & Barak, A. (2012). Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior*, *28*(2), 434–443.

Lederman, R., & Johnston, R. B. (2011). Decision support or support for situated choice: Lessons for system design from effective manual systems. *European Journal of Information Systems*, *20*(5), 510–528.

Ledochowski, L., Ruedl, G., Taylor, A. H., & Kopp, M. (2015). Acute effects of brisk walking on sugary snack cravings in overweight people, affect and responses to a manipulated stress situation and to a sugary snack cue: A crossover study. *PLoS ONE*, *10*(3).

Ledoux, T., Nguyen, A. S., Bakos-Block, C., & Bordnick, P. (2013). Using virtual reality to study food cravings. *Appetite*, *71*, 396–402.

Legaki, N.-Z., Xi, N., Hamari, J., Karpouzis, K., & Assimakopoulos, V. (2020). The effect of challenge-based gamification on learning: An experiment in the context of statistics education. *International Journal of Human-Computer Studies*, 102496.

Lehman, B., D'Mello, S., Strain, A., Mills, C., Gross, M., Dobbins, A., Wallace, P., Millis, K., & Graesser, A. (2013). Inducing and tracking confusion with contradictions during complex learning. *International Journal of Artificial Intelligence in Education*, *22*(1), 85–105.

Leite, I., Martinho, C., & Paiva, A. (2013). Social robots for long-term interaction: A survey. *International Journal of Social Robotics*, *5*(2), 291–308.

Leong, P. (2011). Role of social presence and cognitive absorption in online learning environments. *Distance Education*, *32*(1), 5–28.

Lewicki, R. J., & Bunker, B. B. (1995). Trust in relationships: A model of development and decline. *Conflict, cooperation, and justice: Essays inspired by the work of morton deutsch* (pp. 133–173). Jossey-Bass/Wiley.

Liao, G.-Y., Pham, T. T. L., Cheng, T., & Teng, C.-I. (2020). Impacts of real-world need satisfaction on online gamer loyalty: Perspective of self-affirmation theory. *Computers in Human Behavior*, *103*, 91–100.

Lichtenberg, S., Lembcke, T.-B., Brening, M., Brendel, A. B., & Trang, S. (2020). Can gamification lead to increase paid crowdworkers output? *Proceedings of the 15th International Conference on Business Informatics*.

Lin, H.-F. (2007). Effects of extrinsic and intrinsic motivation on employee knowledge sharing intentions. *Journal of Information Science*, *33*(2), 135–149.

Lin, J. C.-C., & Lu, H. (2000). Towards an understanding of the behavioural intention to use a web site. *International Journal of Information Management*, *20*(3), 197–208.

Lin, Y., Wang, G., & Suh, A. (2020). Exploring the effects of immersive virtual reality on learning outcomes: A two-path model. In D. D. Schmorrow & C. M. Fidopiastis (Eds.), *Augmented Cognition. Human Cognition and Behavior* (pp. 86–105). Springer International Publishing.

Lister, C., West, J. H., Cannon, B., Sax, T., & Brodegard, D. (2014). Just a fad? Gamification in health and fitness apps. *JMIR Serious Games*, *2*(2).

Liu, B. (2010). Sentiment analysis and subjectivity. *Handbook of natural language processing* (p. 38). Chapman; Hall.

Liu, D., Santhanam, R., & Webster, J. (2017). Toward meaningful engagement: A framework for design and research of gamified information systems. *MIS Quarterly*, *41*(4), 1011–1034.

Liu, N., & Yu, R. (2018). Determining effects of virtually and physically present co-actor in evoking social facilitation. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *28*(5), 260–267.

Liu, Y., Alexandrova, T., & Nakajima, T. (2011). Gamifying intelligent environments. *Proceedings of the 2011 International ACM Workshop on Ubiquitous Meta User Interfaces - Ubi-MUI '11*, 7.

Lobato, E. J., Wiltshire, T. J., & Fiore, S. M. (2013). A dual-process approach to understanding human-robot interaction. *Proceedings of the Human Factors and Ergonomics Society*, 1263–1267.

Lohani, M., Stokes, C., McCoy, M., Bailey, C. A., & Rivers, S. E. (2016). Social interaction moderates human-robot trust-reliance relationship and improves stress coping. *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 471–472.

Lohse, M., Hegel, F., & Wrede, B. (2008). Domestic applications for social robots - an online survey on the influence of appearance and capabilities. *Journal of Physical Agents*, *2*(2), 21–32.

Lombriser, P., Dalpiaz, F., Lucassen, G., & Brinkkemper, S. (2016). Gamified requirements engineering: Model and experimentation. In M. Daneva & O. Pastor (Eds.), *Requirements engineering: Foundation for software quality* (pp. 171–187). Springer International Publishing.

Lowry, P., Gaskin, J., Twyman, N., Hammer, B., & Roberts, T. (2013). Taking 'fun and games' seriously: Proposing the hedonic-motivation system adoption model (HMSAM). *Journal of the Association for Information Systems*, *14*(11), 617–671.

Lowry, P. B., Moody, G. D., & Chatterjee, S. (2017). Using IT design to prevent cyberbullying. *Journal of Management Information Systems*, *34*(3), 863–901.

Lowry, P. B., Zhang, J., Moody, G. D., Chatterjee, S., Wang, C., & Wu, T. (2019). An integrative theory addressing cyberharassment in the light of technology-based opportunism. *Journal of Management Information Systems*, *36*(4), 1142–1178.

Lowry, P. B., Zhang, J., Wang, C., & Siponen, M. (2016). Why do adults engage in cyberbullying on social media? an integration of online disinhibition and deindividuation effects with the social structure and social learning model. *Information Systems Research*, *27*(4), 962–986.

Lu, Y., Papagiannidis, S., & Alamanos, E. (2019). Exploring the emotional antecedents and outcomes of technology acceptance. *Computers in Human Behavior*, *90*, 153–169.

Luhmann, N. (1979). *Trust and power*. Wiley.

Lumsden, J., Edwards, E. A., Lawrence, N. S., Coyle, D., & Munafò, M. R. (2016). Gamification of cognitive assessment and cognitive training: A systematic review of applications and efficacy. *JMIR Serious Games*, *4*(2).

MacDorman, K. F., & Chattopadhyay, D. (2016). Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition*, *146*, 190–205.

MacDorman, K. F., & Chattopadhyay, D. (2017). Categorization-based stranger avoidance does not explain the uncanny valley effect. *Cognition*, *161*, 132–135.

MacDorman, K. F., Green, R. D., Ho, C.-C., & Koch, C. T. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior*, *25*(3), 695–710.

Machulska, A. (2021). Approach bias retraining through virtual reality in smokers willing to quit smoking: A randomized-controlled study. *Behaviour Research and Therapy*, 13.

Machulska, A., Zlomuzica, A., Adolph, D., Rinck, M., & Margraf, J. (2015). A cigarette a day keeps the goodies away: Smokers show automatic approach tendencies for smoking—but not for food-related stimuli. *PLOS ONE*, *10*(2), e0116464.

Machulska, A., Zlomuzica, A., Rinck, M., Assion, H.-J., & Margraf, J. (2016). Approach bias modification in inpatient psychiatric smokers. *Journal of Psychiatric Research*, *76*, 44–51.

MacKinnon, D. P., & Fairchild, A. J. (2009). Current directions in mediation analysis. *Current Directions in Psychological Science*, *18*(1), 16–20.

Magni, M., Paolino, C., Cappetta, R., & Proserpio, L. (2013). Diving too deep: How cognitive absorption and group learning behavior affect individual learning. *Academy of Management Learning & Education*, *12*(1), 51–69.

Maister, L., Slater, M., Sanchez-Vives, M. V., & Tsakiris, M. (2015). Changing bodies changes minds: Owning another body affects social cognition. *Trends in Cognitive Sciences*, *19*(1), 6–12.

MarketsandMarkets. (2019). *Gamification in education market by software & services - 2023 | MarketsandMarkets*. Retrieved March 5, 2021, from https://www.marketsandmarkets.com/Market-Reports/gamification-education-market-10910763.html

MarketsandMarkets. (2020). *Gamification market future growth, trends and analysis*. Retrieved March 5, 2021, from https://www.marketsandmarkets.com/Market-Reports/gamification-market-991.html

Martín-Gutiérrez, J., Mora, C. E., Añorbe-Díaz, B., & González-Marrero, A. (2017). Virtual technologies trends in education. *EURASIA Journal of Mathematics, Science and Technology Education*, *13*(1).

Martini, M. C., Gonzalez, C. A., & Wiese, E. (2016). Seeing minds in others – Can agents with robotic appearance have human-like preferences? (M. A. Pavlova, Ed.). *PLOS ONE*, *11*(1), 1–23.

Martocchio, J. J., & Webster, J. (1992). Effects of feedback and cognitive playfulness on performance in microcomputer software training. *Personnel Psychology*, *45*(3), 553–578.

Maselli, A., & Slater, M. (2013). The building blocks of the full body ownership illusion. *Frontiers in Human Neuroscience*, *7*.

Mathur, M. B., & Reichling, D. B. (2009). An uncanny game of trust: Social trustworthiness of robots inferred from subtle anthropomorphic facial cues. *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction - HRI '09*, 313–314.

Mathur, M. B., & Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the uncanny valley. *Cognition*, *146*, 22–32.

Mathur, M. B., Reichling, D. B., Lunardini, F., Geminiani, A., Antonietti, A., Ruijten, P. A., Levitan, C. A., Nave, G., Manfredi, D., Bessette-Symons, B., Szuts, A., & Aczel, B. (2020). Uncanny but not confusing: Multisite study of perceptual category confusion in the uncanny valley. *Computers in Human Behavior*, *103*, 21–30.

May, J., Andrade, J., Kavanagh, D. J., Feeney, G. F. X., Gullo, M. J., Statham, D. J., Skorka-Brown, J., Connolly, J. M., Cassimatis, M., Young, R. M., & Connor, J. P. (2014). The craving experience questionnaire: A brief, theory-based measure of consummatory desire and craving. *Addiction*, *109*(5), 728–735.

McCord, J. (1988). Parental behavior in the cycle of aggression. *Psychiatry*, *51*(1), 14–23.

McKnight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems*, *2*(2), 12:1–12:25.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002a). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research*, *13*(3), 334–359.

McKnight, H., Choudhury, V., & Kacmar, C. (2002b). The impact of initial consumer trust on intentions to transact with a web site: A trust building model. *The Journal of Strategic Information Systems*, *11*(3), 297–323.

Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2017). Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior*, *71*, 525–534.

Metzinger, T. K. (2013). Why are dreams interesting for philosophers? The example of minimal phenomenal selfhood, plus an agenda for future research. *Frontiers in Psychology*, *4*.

Michael, D., & Chen, S. (2005). *Serious games: Games that educate, train, and inform.* Muska & Lipman/Premier-Trade.

Mitchell, R., Schuster, L., & Drennan, J. (2017). Understanding how gamification influences behaviour in social marketing. *Australasian Marketing Journal (AMJ)*, *25*(1), 12–19.

Miura, N., Sugiura, M., Takahashi, M., Miyamoto, A., & Kawashima, R. (2009). The effect of emotional valence and body structure on emotional empathy to humanoid robot: An fMRI study. *NeuroImage*, *47*, 39–41.

Molenberghs, P., & Louis, W. R. (2018). Insights from fMRI studies into ingroup bias. *Frontiers in Psychology*, *9*.

Moon, J. Y., & Sproull, L. S. (2008). The role of feedback in managing the internet-based volunteer work force. *Information Systems Research*, *19*(4), 494–515.

Mori, M. (1970). The uncanny valley. *Energy*, *7*(4), 33–35.

Morrison, S., Decety, J., & Molenberghs, P. (2012). The neuroscience of group membership. *Neuropsychologia*, *50*(8), 2114–2120.

Morschheuser, B., Hamari, J., & Maedche, A. (2019). Cooperation or competition – When do people contribute more? A field experiment on gamification of crowdsourcing. *International Journal of Human-Computer Studies*, *127*, 7–24.

Morschheuser, B., Henzi, C., & Alt, R. (2015). Increasing intranet usage through gamification – Insights from an experiment in the banking industry. *2015 48th Hawaii International Conference on System Sciences*, 635–642.

Morschheuser, B., Maedche, A., & Walter, D. (2017). Designing cooperative gamification: Conceptualization and prototypical implementation. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17*, 2410–2421.

Mueller, M., & Heger, O. (2018). Health at any cost? Investigating ethical dimensions and potential conflicts of an ambulatory therapeutic assistance system through value sensitive design. *Proceedings of the 39th International Conference on Information Systems (ICIS), San Francisco, USA*.

Mühlberger, A., Sperber, M., Wieser, M., & Pauli, P. (2008). A virtual reality behavior avoidance test (VR-BAT) for the assessment of spider phobia. *Journal of Cyber Therapy and Rehabilitation*, *1*, 147–158.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, *20*(8), 750–756.

Muriel, D., & Crawford, G. (2018). *Video games as culture: Considering the role and importance of video games in contemporary society*. Routledge.

Nah, F. F.-H., Eschenbrenner, B., & DeWester, D. (2011). Enhancing brand equity through flow and telepresence: A comparison of 2D and 3D virtual worlds. *MIS Quarterly*, *35*(3), 731–747.

Nakauchi, Y., & Simmons, R. (2002). A social robot that stands in line. *Autonomous Robots*, *12*(3), 313–324.

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 72–78.

NCCPE. (2015). *Future challenges*. Retrieved January 16, 2017, from https://www.publicengagement.ac.uk/explore-it/culture-change/future-challenges

Neimeijer, R. A. M., Roefs, A., Ostafin, B. D., & de Jong, P. J. (2017). Automatic approach tendencies toward high and low caloric food in restrained eaters: Influence of task-relevance and mood. *Frontiers in Psychology*, *8*.

Neto, J. A. M., Yokoyama, K. M., & Becker, K. (2017). Studying toxic behavior influence and player chat in an online video game. *Proceedings of the International Conference on Web Intelligence*, 26–33.

Newzoo. (2019a). *Most popular core PC games | global*. Retrieved April 24, 2019, from https://%20newzoo.com/insights/rankings/top-20-core-pc-games/

Newzoo. (2019b). *NewNewzoo's 2018 report: Insights into the $137.9 billion global games marketzoo*. Retrieved April 25, 2019, from https://newzoo.com/insights/articles/newzoos-2018-report-insightsinto-%20the-137-9-billion-global-games-market/

Niehaves, B., & Ortbach, K. (2016). The inner and the outer model in explanatory design theory: The case of designing electronic feedback systems. *European Journal of Information Systems*, *25*(4), 303–316.

Nier, J. A., Gaertner, S. L., Dovidio, J. F., Banker, B. S., Ward, C. M., & Rust, M. C. (2001). Changing interracial evaluations and behavior: The effects of a common group identity. *Group Processes & Intergroup Relations*, *4*(4), 299–316.

Nippert-Eng, C. E. (1996). *Home and work: Negotiating boundaries through everyday life*. University of Chicago Press.

Nissen, A., & Jahn, K. (2021). Between anthropomorphism, trust, and the uncanny valley: A dual-processing perspective on perceived trustworthiness and its mediating effects on use intentions of social robots. *Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS-54), Maui, Hawaii, USA*, 360–369.

Nomura, T. (2017). Robots and Gender. *Gender and the Genome*, *1*(1), 18–26.

Oh, S. Y., Bailenson, J., Weisz, E., & Zaki, J. (2016). Virtually old: Embodied perspective taking and the reduction of ageism under threat. *Computers in Human Behavior*, *60*, 398–410.

Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in facebook and its application to e-learning. *Computers in Human Behavior*, *31*, 527–541.

Osimo, S. A., Pizarro, R., Spanlang, B., & Slater, M. (2015). Conversations between self and self as Sigmund Freud: A virtual body ownership paradigm for self counselling. *Scientific Reports*, *5*, 13899.

Ott, M., & Freina, L. (2015). A literature review on immersive virtual reality in education: State of the art and perspectives. *Conference proceedings of »eLearning and Software for Education« (eLSE)*, 133–141.

Owen, R., Stigloe, J., Macnaghten, P., Gorman, M., Fisher, E., & Guston, D. (2013). A framework for responsible innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible innovation: Managing the responsible emergence of science and innovation in society* (pp. 27–50). Wiley.

Oyelere, S. S., Bouali, N., Kaliisa, R., Obaido, G., Yunusa, A. A., & Jimoh, E. R. (2020). Exploring the trends of educational virtual reality games: A systematic review of empirical studies. *Smart Learning Environments*, *7*(1), 1–22.

Pabian, S., & Vandebosch, H. (2014). Using the theory of planned behaviour to understand cyberbullying: The importance of beliefs for developing interventions. *European Journal of Developmental Psychology*, *11*(4), 463–477.

Pacherie, E. (2007). The sense of control and the sense of agency. *PSYCHE: An Interdisciplinary Journal of Research On Consciousness*, *13*(1), 1–30.

Paetzel, M., Perugia, G., & Castellano, G. (2020). The persistence of first impressions: The effect of repeated interactions on the perception of a social robot. *ACM/IEEE International Conference on Human-Robot Interaction*, 73–82.

Park, S., & Catrambone, R. (2007). Social Facilitation Effects of Virtual Humans. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *49*(6), 1054–1060.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first date of life? *Science*, *296*(5571), 1321–1323.

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779–787.

Pelletier, K. L. (2010). Leader toxicity: An empirical investigation of toxic behavior and rhetoric. *Leadership*, *6*(4), 373–389.

Persky, S. (2011). Application of virtual reality methods to obesity prevention and management research. *Journal of Diabetes Science and Technology*, *5*(2), 333–339.

Petkova, V. I., & Ehrsson, H. H. (2008). If I were you: Perceptual illusion of body swapping (J. Harris, Ed.). *PLoS ONE*, *3*(12), e3832.

Petkova, V. I., Khoshnevis, M., & Ehrsson, H. H. (2011). The perspective matters! Multisensory integration in ego-centric reference frames determines full-body ownership. *Frontiers in Psychology*, *2*.

Pfeuffer, N., Benlian, A., Gimpel, H., & Hinz, O. (2019). Anthropomorphic information systems. *Business and Information Systems Engineering*, *61*(4), 523–533.

Picard, R. W. (1997). *Affective computing*. MIT Press.

Picard, R. W. (2003). Affective computing: Challenges. *International Journal of Human-Computer Studies*, *59*(1), 55–64.

Picard, R. W. (2015). The promise of affective computing. In R. A. Calvo, S. K. D'Mello, J. Gratch, & A. Kappas (Eds.), *The oxford handbook of affective computing* (pp. 11–20). Oxford University Press.

Pietraszewski, D., Cosmides, L., & Tooby, J. (2014). The content of our cooperation, not the color of our skin: An alliance detection system regulates categorization by coalition and race, but not sex. *PLoS ONE*, *9*(2), e88534.

Poelman, M. P., Gillebaart, M., Schlinkert, C., Dijkstra, S. C., Derksen, E., Mensink, F., Hermans, R. C. J., Aardening, P., de Ridder, D., & de Vet, E. (2021). Eating behavior and food purchases during the COVID-19 lockdown: A cross-sectional study among adults in the Netherlands. *Appetite*, *157*, 105002.

Poeschl, S., & Doering, N. (2015). Measuring co-presence and social presence in virtual environments – Psychometric construction of a German scale for a fear of public speaking scenario. *Studies in Health Technology and Informatics*, 58–63.

Polites, G. L., Serrano, C., Thatcher, J. B., & Matthews, K. (2018). Understanding social networking site (SNS) identity from a dual systems perspective: An investigation of the dark side of SNS use. *European Journal of Information Systems*, *27*(5), 600–621.

Pollatos, O., Herbert, B. M., Matthias, E., & Schandry, R. (2007). Heart rate response after emotional picture presentation is modulated by interoceptive awareness. *International Journal of Psychophysiology*, *63*(1), 117–124.

Poondej, C., & Lerdpornkulrat, T. (2016). The development of gamified learning activities to increase student engagement in learning. *Australian Educational Computing*, *31*(2).

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of General Psychology*, *14*(2), 154–166.

RankedKings. (2019). *How many people play league of legends in 2019?* Retrieved April 4, 2021, from https://rankedkings.com/blog/how-many-people-play-league-of-legends

Reed, G. (2004). Toxic leadership. *Military Review*, *84*, 67–71.

Regan, D. T., & Totten, J. (1975). Empathy and attribution: Turning observers into actors. *Journal of Personality and Social Psychology*, *32*(5), 850–856.

Reiner, M., & Hecht, D. (2009). Behavioral indications of object-presence in haptic virtual environments. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *12*, 183–186.

Reiss, S. (2004). Multifaceted nature of intrinsic motivation: The theory of 16 basic desires. *Review of General Psychology*, *8*(3), 179–193.

Reychav, I., & Wu, D. (2015). Are your users actively involved? A cognitive absorption perspective in mobile training. *Computers in Human Behavior*, *44*, 335–346.

Reynolds, C., & Picard, R. (2004). Affective sensors, privacy, and ethical contracts. *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, 1103–1106.

Reyt, J.-N., & Wiesenfeld, B. M. (2015). Seeing the forest for the trees: Exploratory learning, mobile technology, and knowledge workers' role integration behaviors. *Academy of Management Journal*, *58*, 739–762.

Ridder, J. M. M. v. d., McGaghie, W. C., Stokking, K. M., & Cate, O. T. J. t. (2015). Variables that affect the process and outcome of feedback, relevant for medical training: A meta-review. *Medical Education*, *49*(7), 658–673.

Riedl, R., Hubert, M., & Kenning, P. (2010). Are there neural gender differences in online trust? An fMRI study on the perceived trustworthiness of eBay offers. *MIS Quarterly*, *34*(2), 397–428.

Riedl, R., Mohr, P., Kenning, P., Davis, F., & Heekeren, H. (2014). Trusting humans and avatars: A brain imaging study based on evolution theory. *Journal of Management Information Systems*, *30*(4), 83–114.

Rigby, C. S., & Ryan, R. M. (2011). *Glued to games: How video games draw us in and hold us spellbound*. ABC.

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(2), 105–120.

Riva, G. (2009). Is presence a technology issue? Some insights from cognitive sciences. *Virtual Reality*, *13*(3), 159–169.

RoadToVR. (201828). *Volkswagen group to train 10,000 employees in VR this year*. Retrieved March 5, 2018, from https://www.roadtovr.com/volkswagen-group-to-train-10000-employees-in-vr-in-2018/

Robinson, D., & Bellotti, V. (2013). A preliminary taxonomy of gamification elements for varying anticipated commitment. *Proc. ACM CHI 2013 Workshop on Designing Gamification. Creating Gameful and Playful Experiences, Paris, France*.

Roca, J. C., & Gagné, M. (2008). Understanding e-learning continuance intention in the workplace: A self-determination theory perspective. *Computers in Human Behavior*, *24*(4), 1585–1604.

Rode, H. (2016). To share or not to share: The effects of extrinsic and intrinsic motivations on knowledge-sharing in enterprise social media platforms. *Journal of Information Technology*, *31*(2), 152–165.

Roelofs, K., Minelli, A., Mars, R. B., van Peer, J., & Toni, I. (2009). On the neural control of social emotional behavior. *Social Cognitive and Affective Neuroscience*, *4*(1), 50–58.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, Articles*, *48*(2), 1–36.

Rothbard, N. P., Phillips, K. W., & Dumas, T. L. (2005). Managing multiple roles: Work-family policies and individuals' desires for segmentation. *Organization Science*, *16*(3), 243–258.

Rotolo, D., Hicks, D., & Martin, B. R. (2015). What is an emerging technology? *Research Policy*, *44*(10), 1827–1843.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, *23*(3), 393–404.

Runions, K. C., & Bak, M. (2015). Online moral disengagement, cyberbullying, and cyber-aggression. *Cyberpsychology, Behavior, and Social Networking*, *18*(7), 400–405.

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, 68–78.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press.

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61*, 101860.

Ryan, R. M., Rigby, C. S., & Przybylski, A. K. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion*, *30*(4), 344–360.

Sailer, M. (2016). *Die Wirkung von Gamification auf Motivation und Leistung: Empirische Studien im Kontext manueller Arbeitsprozesse [engl. The effect of gamification on motivation and performance: Empirical studies in the context of manual working processes]*. Springer Fachmedien.

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, *69*, 371–380.

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review*, *32*(1), 77–112.

Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015). Would you trust a (faulty) robot? Effects of error, task type and personality on human-robot cooperation and trust. *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 1–8.

Sanchez-Vives, M. V., Spanlang, B., Frisoli, A., Bergamasco, M., & Slater, M. (2010). Virtual hand illusion induced by visuomotor correlations (M. W. Greenlee, Ed.). *PLoS ONE*, *5*(4), e10381.

Sanders, T., Oleson, K. E., Billings, D. R., Chen, J. Y. C., & Hancock, P. A. (2011). A model of human-robot trust: Theoretical model development. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *55*(1), 1432–1436.

Sarmany-Schuller, I. (2010). Decision making under time pressure in regard to preferred cognitive style (analytical-intuitive) and study orientation. *Studia Psychologia*, *52*(4), 285–290.

Saygin, A.P., Chaminade, T., & Ishiguro, H. (2010). The perception of humans and robots: Uncanny hills in parietal cortex. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society (pp. 2716-2720). Austin, TX: Cognitive Science Society.*

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-offit measures. *Methods of Psychological Research*, *8*, 23–74.

Schindler, S., Zell, E., Botsch, M., & Kissler, J. (2017). Differential effects of face-realism and emotion on event-related brain potentials and their implications for the uncanny valley theory. *Scientific Reports*, *7*(1), 45003.

Schöbel, S., & Janson, A. (2018). Is it all about Having Fun? - Developing a Taxonomy to Gamify Information Systems. *Proceedings of the 26th European Conference on Information Systems (ECIS), Portsmouth, UK.*

Schöbel, S., Janson, A., Jahn, K., Kordyaka, B., Turetken, O., Djafarova, N., Saqr, M., Wu, D., Söllner, M., Adam, M., Heiberg Gad, P., Wesseloh, H., Leimeister, J. M., Schöbel, S., Janson, A., Jahn, K., Kordyaka, B., Turetken, O., Djafarova, N., … Leimeister, J. (2020). A research agenda for the why, what, and how of gamification designs results on an ECIS 2019 panel. *Communications of the Association for Information Systems*, *46*(1), 706–721.

Schöbel, S., Söllner, M., & Leimeister, J. M. (2017). The agony of choice – analyzing user preferences regarding gamification elements in learning management systems. *Proceedings of the 37th International Conference on Information Systems (ICIS), Dublin, Ireland.*

Schoenmakers, T. M., de Bruin, M., Lux, I. F. M., Goertz, A. G., Van Kerkhof, D. H. A. T., & Wiers, R. W. (2010). Clinical effectiveness of attentional bias modification training in abstinent alcoholic patients. *Drug and Alcohol Dependence*, *109*(1–3), 30–36.

Schroeder, P. A., Lohmann, J., Butz, M. V., & Plewnia, C. (2016). Behavioral bias for food reflected in hand movements: A preliminary study with healthy subjects. *Cyberpsychology, Behavior, and Social Networking*, *19*(2), 120–126.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments*, *10*(3), 266–281.

Schultze, U. (2010). Embodiment and presence in virtual worlds: A review. *Journal of Information Technology*, *25*(4), 434–449.

Schultze, U. (2014). Performing embodied identity in virtual worlds. *European Journal of Information Systems*, *23*(1), 84–95.

Schultze, U., & Orlikowski, W. J. (2010). Virtual worlds: A performative perspective on globally distributed, immersive work. *Information Systems Research*, *21*(4), 810–821.

Schumacher, S. E., Kemps, E., & Tiggemann, M. (2016). Bias modification training can alter approach bias and chocolate consumption. *Appetite*, *96*, 219–224.

Schwind, V., Knierim, P., Tasci, C., Franczak, P., Haas, N., & Henze, N. (2017). These are not my hands! Effect of gender on the perception of avatar hands in virtual reality. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1577–1582.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *74*, 14–31.

Seibt, B., Häfner, M., & Deutsch, R. (2007). Prepared to eat: How immediate affective and motivational responses to food cues are influenced by food deprivation. *European Journal of Social Psychology*, *37*(2), 359–379.

Seo, Y., Dolan, R., & Buchanan-Oliver, M. (2019). Playing games: Advancing research on online and mobile gaming consumption. *Internet Research*, *29*(2), 289–292.

Sesko, A. K., & Biernat, M. (2010). Prototypes of race and gender: The invisibility of black women. *Journal of Experimental Social Psychology*, *46*(2), 356–360.

Seyama, J., & Nagayama, R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *PRESENCE: Teleoperators and Virtual Environments*, *16*, 337–351.

Seymour, M., Riemer, K., & Kay, J. (2017). Interactive realistic digital avatars - revisiting the uncanny valley. *Proceedings of the 50th Hawaii International Conference on System Sciences (2017)*, 547–556.

Shah, J. N., & Castell, S. (2011). *Public attitudes to science 2011*. Ipsos MORI Social Research Institute.

Shi, L., Cristea, A. I., Hadzidedic, S., & Dervishalidovic, N. (2014). Contextual gamification of social interaction – towards increasing motivation in social e-learning. In E. Popescu, R. W. H. Lau, K. Pata, H. Leung, & M. Laanpere (Eds.), *Advances in web-based learning – ICWL 2014* (pp. 116–122). Springer International Publishing.

Shores, K. B., He, Y., Swanenburg, K. L., Kraut, R., & Riedl, J. (2014). The identification of deviance and its impact on retention in a multiplayer game. *Proceedings* *of the 17th ACM conference on Computer supported cooperative work & social computing*, 1356–1365.

Siegel, M., Breazeal, C., & Norton, M. I. (2009). Persuasive Robotics: The influence of robot gender on human behavior. *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2563–2568.

Simpson, D. D., & Joe, G. W. (2004). A longitudinal evaluation of treatment engagement and recovery stages. *Journal of Substance Abuse Treatment*, *27*(2), 89–97.

Skyes, K., & Macnaghten, P. (2013). Responsible innovation – opening up dialogue and debate. *Responsible innovation - managing the responsible emergence of science and innovation in society* (pp. 85–107). Wiley.

Slater, M., Spanlang, B., Sanchez-Vives, M. V., & Blanke, O. (2010). First person experience of body transfer in virtual reality. *PloS One*, *5*(5), e10564.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, *6*(6), 603–616.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1), 3–22.

Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, *6*(4), 342–353.

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review*, *8*(2), 203–220.

Smits, M. H. S. B., Boon, J., Sluijsmans, D. M. A., & Gog, T. v. (2008). Content and timing of feedback in a web-based learning environment: Effects on learning as a function of prior knowledge. *Interactive Learning Environments*, *16*(2), 183–193.

Snyder, A. L., Anderson-Hanley, C., & Arciero, P. J. (2012). Virtual and Live Social Facilitation while Exergaming: Competitiveness moderates exercise intensity. *Journal of Sport and Exercise Psychology*, *34*(2), 252–259.

Sørebø, Ø., Halvari, H., Gulli, V. F., & Kristiansen, R. (2009). The role of self-determination theory in explaining teachers' motivation to continue to use e-learning technology. *Computers & Education*, *53*(4), 1177–1187.

Soror, A. A., Hammer, B. I., Steelman, Z. R., Davis, F. D., & Limayem, M. M. (2015). Good habits gone bad: Explaining negative consequences associated with the use of mobile phones from a dual-systems perspective. *Information Systems Journal*, *25*(4), 403–427.

Sourina, O., & Liu, Y. (2013). EEG-enabled affective applications. *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 707–708.

Soutter, A. R. B., & Hitchens, M. (2016). The relationship between character identification and flow state within video games. *Computers in Human Behavior*, *55*, 1030–1038.

Spinuzzi, C. (2005). The methodology of participatory design. *Technical Communication*, *52*(2), 163–174.

Stansbury, J. A., & Earnest, D. R. (2017). Meaningful gamification in an industrial/organizational psychology course. *Teaching of Psychology*, *44*(1), 38–45.

Statista. (2019). *LoL global revenue 2018*. Retrieved July 28, 2019, from https://www.statista.com/%20statistics/806975/lol-revenue/

Statista. (2020). *Number of active video gamers worldwide from 2015 to 2023*. Retrieved January 18, 2021, from https://www.statista.com/statistics/748044/number-video-gamers-world/

Steiner, C. (1997). *Achieving emotional literacy*. Bloomsburry Publishing.

Stevens, B., & Jerrams-Smith, J. (2001). The sense of object-presence with projection-augmented models. In S. Brewster & R. Murray-Smith (Eds.), *Haptic Human-Computer Interaction. Haptic HCI 2000. Lecture Notes in Computer Science*. Springer.

Stins, J. F., Roelofs, K., Villan, J., Kooijman, K., Hagenaars, M. A., & Beek, P. J. (2011). Walk to me when i smile, step back when i'm angry: Emotional faces modulate whole-body approach-avoidance behaviors. *Experimental Brain Research*, *212*(4), 603–611.

Stock, R., Merkle, M., Eidens, D., Hannig, M., Heineck, P., Nguyen, M. A., & Völker, J. (2019). When robots enter our workplace: Understanding employee trust in assistive robots. *40th International Conference on Information Systems (ICIS)*, 1–9.

Stoltz, D. S., & Lizardo, O. (2018). Deliberate trust and intuitive faith: A dual-process model of reliance. *Journal for the Theory of Social Behaviour*, *48*(2), 230–250.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3), 220–247.

Strait, M., Canning, C., & Scheutz, M. (2014). Investigating the effects of robot communication strategies in advice-giving situations based on robot appearance, interaction modality and distance. *ACM/IEEE International Conference on Human-Robot Interaction*, 479–486.

Strait, M., Urry, H. L., & Muentener, P. (2019). Children's responding to human-like agents reflects an uncanny valley. *ACM/IEEE International Conference on Human-Robot Interaction*, *2019*, 506–515.

Strait, M. K., Floerke, V. A., Ju, W., Maddox, K., Remedios, J. D., Jung, M. F., & Urry, H. L. (2017). Understanding the uncanny: Both atypical features and category ambiguity provoke aversion toward humanlike robots. *Frontiers in Psychology*, *8*.

Stratou, G., Morency, L.-P., DeVault, D., Hartholt, A., Fast, E., Lhommet, M., Lucas, G., Morbini, F., Georgila, K., Scherer, S., Gratch, J., Marsella, S., Traum, D., & Rizzo, A. (2015). A demonstration of the perception system in SimSensei, a virtual human application for healthcare interviews. *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, 787–789.

Sturges, J. (2012). Crafting a balance between work and home. *Human Relations*, *65*(12), 1539–1559.

Sugano, N., Kato, H., & Tachibana, K. (2003). The effects of shadow representation of virtual objects in augmented reality. *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.*, 76–83.

Suh, K.-S., Kim, H., & Suh, E. K. (2011). What if your avatar looks like you? Dual-congruity perspectives for avatar use. *MIS Quarterly*, *35*(3), 711–729.

Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior*, *7*(3), 321–326.

Sylaiou, S., Mania, K., Karoulis, A., & White, M. (2010). Exploring the relationship between presence and enjoyment in a virtual museum. *International Journal of Human-Computer Studies*, *68*(5), 243–253.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of inter group behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 33–47). Brooks/Cole.

Tal, A., & Wansink, B. (2011). Turning virtual reality into reality: A checklist to ensure virtual reality studies of eating behavior and physical activity parallel the real world. *Journal of Diabetes Science and Technology*, *5*(2), 239–244.

Tarafdar, M., Gupta, A., & Turel, O. (2013). The dark side of information technology use. *Information Systems Journal*, *23*(3), 269–275.

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, *36*(7), 778–793.

Taylor, T. L. (2002). Living digitally: Embodiment in virtual worlds. In R. Schroeder (Ed.), *The social life of avatars* (pp. 40–62). Springer.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth and happiness* (1st ed.). Penguin Books.

Thiebes, S., Lins, S., & Basten, D. (2014). Gamifying information systems - A synthesis of gamification mechanics and dynamics. *Proceedings of the European Conference on Information Systems (ECIS) 2014, Tel Aviv, Israel.*

Tieri, G., Tidoni, E., Pavone, E. F., & Aglioti, S. M. (2015). Body visual discontinuity affects feeling of ownership and skin conductance responses. *Scientific Reports*, *5*.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis. *Journal of Statistical Software*, *59*(5).

Triberti, S., & Riva, G. (2016). Being present in action: A theoretical model about the "interlocking" between intentions and environmental affordances. *Frontiers in Psychology*, *6*.

Tsai, A. G., Williamson, D. F., & Glick, H. A. (2011). Direct medical cost of overweight and obesity in the USA: A quantitative systematic review. *Obesity Reviews*, *12*(1), 50–61.

Udris, R. (2014). Cyberbullying among high school students in Japan: Development and validation of the online disinhibition scale. *Computers in Human Behavior*, *41*, 253–261.

Urquhart, C., Lehmann, H., & Myers, M. D. (2010). Putting the 'theory' back into grounded theory: Guidelines for grounded theory studies in information systems. *Information Systems Journal*, *20*(4), 357–381.

Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting.* Oxford University Press.

van den Akker, K., Schyns, G., & Jansen, A. (2018). Learned overeating: Applying principles of Pavlovian conditioning to explain and treat overeating. *Current Addiction Reports*, *5*(2), 223–231.

van den Hooff, B., & De Leeuw van Weenen, F. d. L. (2004). Committed to share: Commitment and CMC use as antecedents of knowledge sharing. *Knowledge and Process Management*, *11*(1), 13–24.

van den Hoven, J. (2013). Value sensitive design and responsible innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible innovation.* Wiley.

van der Heijden. (2004). User acceptance of hedonic information systems. *MIS Quarterly*, *28*(4), 695–704.

van der Land, S. F., Schouten, A. P., Feldberg, F., Huysman, M., & van den Hooff, B. (2015). Does avatar appearance matter? How team visual similarity and member–avatar similarity influence virtual team performance. *Human Communication Research*, *41*(1), 128–153.

Vandenberg, R. J. (2006). Introduction: Statistical and methodological myths and urban legends: Where, pray tell, did they get this idea? *Organizational Research Methods*, *9*(2), 194–201.

van Roy, R., & Zaman, B. (2019). Unravelling the ambivalent motivational power of gamification: A basic psychological needs perspective. *International Journal of Human-Computer Studies*, *127*, 38–50.

Vansteenkiste, M., Niemiec, C. P., & Soenens, B. (2010). The development of the five mini-theories of self-determination theory: An historical overview, emerging trends, and future directions. In T. C. Urdan & S. A. Karabenick (Eds.), *Advances in motivation and achievement* (pp. 105–165). Emerald Group Publishing Limited.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478.

Verified Market Research. (2021). *Virtual Reality in Healthcare*. Retrieved March 31, 2021, from https://www.verifiedmarketresearch.com/product/virtual-reality-in-healthcare-market/

Vescio, T. K., Judd, C. M., & Kwan, V. S. (2004). The crossed-categorization hypothesis: Evidence of reductions in the strength of categorization, but not intergroup bias. *Journal of Experimental Social Psychology*, *40*(4), 478–496.

vom Brocke, J., Hevner, A., Léger, P. M., Walla, P., & Riedl, R. (2020). Advancing a NeuroIS research agenda with four areas of societal contributions. *European Journal of Information Systems*, *29*(1), 9–24.

von Hippel, E. (2009). Democratizing innovation: The evolving phenomenon of user innovation. *International Journal of Innovation Science*, *1*(1), 29–40.

Wang, S., & Noe, R. A. (2010). Knowledge sharing: A review and directions for future research. *Human Resource Management Review*, *20*(2), 115–131.

Wang, Y., & Quadflieg, S. (2014). In our own image? Emotional and neural processing differences when observing human-human vs human-robot interactions. *Social Cognitive and Affective Neuroscience*, *10*(11), 1515–1524.

Warburton, D., Colbourne, L., Gavelin, K., Wilson, R., & Noun, A. (2008). *Deliberative public engagement: Nine principles*. National Consumer Council.

Watanabe, K., & Fukuta, N. (2017). Toward empathic agents for defusing toxic behaviors on team competition games. *2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, 820–824.

Weingarten, H. P., & Elston, D. (1991). Food cravings in a college population. *Appetite*, *17*(3), 167–175.

Wenhui, P., Ruixue, P., Jun, L., Yan, C., & Xiaojie, G. (2009). Application study of the tracking of e-learning behavior using the net-based intelligent robot. *2009 First International Workshop on Education Technology and Computer Science*, *2*, 529–533.

Wiederhold, B. K., Riva, G., & Gutiérrez-Maldonado, J. (2016). Virtual reality in the assessment and treatment of weight-related disorders. *Cyberpsychology, Behavior, and Social Networking*, *19*(2), 67–73.

Wiers, C. E., Gladwin, T. E., Ludwig, V. U., Gröpper, S., Stuke, H., Gawron, C. K., Wiers, R. W., Walter, H., & Bermpohl, F. (2016). Comparing three cognitive biases for alcohol cues in alcohol dependence. *Alcohol and Alcoholism*, *52*(2), 242–248.

Wiers, C. E., Kühn, S., Javadi, A. H., Korucuoglu, O., Wiers, R. W., Walter, H., Gallinat, J., & Bermpohl, F. (2013). Automatic approach bias towards smoking cues is present in smokers but not in ex-smokers. *Psychopharmacology*, *229*(1), 187–197.

Wiers, R. W., Rinck, M., Kordts, R., Houben, K., & Strack, F. (2010). Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction*, *105*(2), 279–287.

Wilsdon, J., Wynne, B., & Stilgoe, J. (2005). *The public value of science*. Demos.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225–240.

World Health Organization. (2018). *Obesity and overweight*. Retrieved May 4, 2019, from https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

Wu, S., Lin, T.-C., & Shih, J.-F. (2017). Examining the antecedents of online disinhibition. *Information Technology & People*, *30*(1), 189–209.

Xi, N., & Hamari, J. (2019). Does gamification satisfy needs? A study on the relationship between gamification features and intrinsic need satisfaction. *International Journal of Information Management*, *46*, 210–221.

Xi, N., & Hamari, J. (2020). Does gamification affect brand engagement and equity? A study in online brand communities. *Journal of Business Research*, *109*, 449–460.

Xiao, B., & Wong, R. Y. M. (2013). Cyber-bullying among university students: An empirical investigation from social cognitive perspective. *International Journal of Business and Information*, *8*(1), 34–69.

Yee, N., Bailenson, J. N., & Ducheneaut, N. (2009). The proteus effect: Implications of transformed digital self-representation on online and offline behavior. *Communication Research*, *36*(2), 285–312.

Yee, N., Ducheneaut, N., & Nelson, L. (2012). Online gaming motivations scale: Development and validation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2803–2806.

Zajonc, R. B. (1965). Social facilitation. *Science*, *149*(3681), 269–274.

Zanbaka, C., Ulinski, A., Goolkasian, P., & Hodges, L. F. (2004). Effects of virtual human presence on task performance. *Proceedings of the International Conference on Artificial Reality and Telexistence*, 174–181.

Zanbaka, C., Ulinski, A., Goolkasian, P., & Hodges, L. F. (2007). Social responses to virtual humans: Implications for future interface design. *Proceedings of the 2007 Conference on Human Factors in Computing Systems (CHI 2007), San Jose, California, USA*, 1561–1570.

Zhang, M. W., Ying, J. B., Song, G., & Ho, R. C. (2018). A review of gamification approaches in commercial cognitive bias modification gaming applications. *Technology and Health Care*, *26*(6), 933–944.

Zhang, X., Guo, X., Lai, K.-h., & Yi, W. (2019). How does online interactional unfairness matter for patient–doctor relationship quality in online health consultation? The contingencies of professional seniority and disease severity. *European Journal of Information Systems*, *28*(3), 336–354.

Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models: Problems and solutions. *Organizational Research Methods*, *12*(4), 695–719.

Złotowski, J., Sumioka, H., Eyssel, F., Nishio, S., Bartneck, C., & Ishiguro, H. (2018). Model of dual anthropomorphism: The relationship between the media equation effect and implicit anthropomorphism. *International Journal of Social Robotics*, *10*(5), 701–714.

Zlotowski, J., Sumioka, H., Nishio, S., Glas, D. F., Bartneck, C., & Ishiguro, H. (2016). Appearance of a robot affects the impact of its behaviour on perceived trustworthiness and empathy. *Paladyn*, *7*(1), 55–66.

Zucker, L. G. (1986). Production of trust: Institutional sources of economic structure, 1840–1920. *Research in Organizational Behavior*, *8*, 53–111.