Research Papers   ECIS 2019 Proceedings

5-15-2019

TO NUDGE OR NOT TO NUDGE: ETHICAL CONSIDERATIONS OF DIGITAL NUDGING BASED ON ITS BEHAVIORAL ECONOMICS ROOTS

Tim-Benjamin Lembcke
University of Goettingen, tim-benjamin.lembcke@uni-goettingen.de

Nils Engelbrecht
University of Goettingen, nils.engelbrecht@wiwi.uni-goettingen.de

Alfred Benedikt Brendel
University of Goettingen, abrende1@gwdg.de

Lutz Kolbe
University of Göttingen, lkolbe@uni-goettingen.de

Follow this and additional works at: https://aisel.aisnet.org/ecis2019_rp

TO NUDGE OR NOT TO NUDGE:
ETHICAL CONSIDERATIONS OF DIGITAL NUDGING BASED ON ITS BEHAVIORAL ECONOMICS ROOTS

Research paper

Lembcke, Tim-Benjamin, University of Goettingen, Goettingen, Germany, tim-benjamin.lembcke@uni-goettingen.de

Engelbrecht, Nils, University of Goettingen, Goettingen, Germany, nils.engelbrecht@uni-goettingen.de

Brendel, Alfred Benedikt, University of Goettingen, Goettingen, Germany, abrende1@uni-goettingen.de

Kolbe, Lutz M., University of Goettingen, Goettingen, Germany, lutz.kolbe@uni-goettingen.de

Abstract

Nudging has become a well renowned concept in policymaking around the globe, supporting active behavior changes like higher retirement saving, healthier nutrition or saving natural resources. Stemming from the behavioral economics (BE), nudge-interventions aim at changing individuals' behaviors without limiting their freedom of choice, exerting coercion or significantly changing economic incentives. As nudging still impairs individuals' autonomy, the concept has evoked heated debates about its ethical acceptability.

Recently, IS scholars have begun to recognize the potential of nudging in digital contexts leading to the notion of digital nudging. The importance of ethical considerations when designing and implementing nudge-interventions has been acknowledged, however, a detailed analysis of these concerns in digital environments is still missing.

In this article, we summarize, transfer and discuss the ethical considerations raised in BE in light of digital contexts. Three important ethical considerations for digital nudges are discussed: (1) Preserving individuals' freedom of choice / autonomy, (2) transparent disclosure of nudges and (3) individual (pro-self) and societal (pro-social) goal-oriented justification of nudging.

As a result, we provide a first ethical landscape for digital nudging in IS, based on its BE roots. We conclude with a provision of several future research avenues.

Keywords: Digital Nudging; Ethics, Justification; Behavioral Economics; Digital Choice Environments; Information Systems Research.

1 Introduction

As digitalization increasingly permeates our everyday life, not only are Information Systems (IS) more actively used but also more decisions are made utilizing IS as decision support mediums. Not only are more decisions supported by IS, but also the scope and significance of digitally-supported decisions rise. Alongside this, more decision support is needed since more information is accessible, possibly causing an information overload or a state of paralysis in regard to using IS in an intended way. The IS community itself has started an interest in linkages between Behavioral Economics (BE) and IS (Goes, 2013), e.g. to better understand different BE phenomena like preference building effects, recommendation systems, and gamification.

Contrary to the predictions made by standard economic theory ("homo oeconomicus"), humans do not always act with full rationality. Instead, each individual action is embedded in an omnipresent and exogenously defined decision context. The specific structure and display of this "choice architecture" can influence the manner in which individuals behave and what decisions they make (Thaler and Sunstein, 2008; Johnson et al., 2012; Münscher et al., 2016; Barton and Grüne-Yanoff, 2015; Szaszi et al., 2018; Thaler et al., 2013). Empirical insights from BE have illustrated that even small and insignificant modifications of choice architectures can strongly impact human decision-making (Münscher et al., 2016; Thaler, 2016). This includes "seemingly irrelevant factors" (Thaler, 2016), like position effects, default rules, framing descriptions, and accounting (Thaler, 2016). Based on these empirical results, a subset of specific behavior change techniques has been developed. The so-called *nudges* aim to be empirically informed, less intrusive, and preserving the freedom of choice. They are therefore argued to be a less intrusive tool for choice architects / designers ("nudge") than alternative choice-limiting interventions like bans, fines, coercion or mandates. As nudges impair individuals' autonomy, the concept has initiated heated debates about the ethical acceptability of nudging (Bovens, 2009; White, 2013).

For several reasons, scholars have begun to recognize the importance of nudging in the context of IS leading to the notion of digital nudging (Weinmann et al., 2016; Schneider et al., 2018; Hummel et al., 2018; Mirsch et al., 2017; Meske and Potthoff, 2017): First, nudging stems from a policy setting context, with most behaviorally-informed instruments and their underlying systematic cognitive boundaries being focused on the "analog world" and therefore cannot per se be embodied in digital choice environments (Benartzi et al., 2017). Second, the increasing digitalization of our society leads to an increased usage of digital decision support systems (DSS), which include devices like smartphones. As context generally alters the specific structure and display of choice architectures and can influence the manner in which individuals behave and what decisions they make (Thaler and Sunstein, 2008; Johnson et al., 2012; Münscher et al., 2016; Barton and Grüne-Yanoff, 2015; Szaszi et al., 2018; Thaler et al., 2013), this new digital context especially has to be taken into consideration. Although ethical considerations for digital nudging have been generally acknowledged (Weinmann et al., 2016; Meske and Potthoff, 2017), specific guidelines or a general debate about digital nudging's ethical acceptability is still missing.

In reply to Meske and Potthoff (2017) and Weinmann et al. (2016), this article argues in favour of starting the important discussion on the ethical foundations of digital nudging in order to extend the current implementation- and process-focused research on digital nudging. So far, digital nudging is understood as "the use of user-interface design elements to guide people's behavior in digital choice environments" (Weinmann et al., 2016). Although we will argue in this article that this definition is too brief as relevant ethical concerns are currently missing, it provides a starting point to the question of how nudging mechanisms can be utilized in a digital environment.

To the best of our knowledge, this article is the first to: (1) offer a clear theoretical linkage to the analog roots of digital nudging, (2) highlight ethical considerations for analog nudges, (3) analyze the influence of (digital) IS on ethical considerations of (analog) nudging, (4) provide the IS community with an advanced notion of important ethical considerations for digital nudging and (5) establish a starting point for future discussion on the ethical usage of digital nudging both for researchers (indicating future research directions) and practitioners (allowing them to better understand ethical considerations of digital nudging in order to design ethically sound IS / DSS).

Our guiding research questions are:

> *RQ1: Which ethical considerations do arise when using nudging mechanisms in digital choice environments (i.e. digital nudging)?*
>
> *RQ2: What potential resolutions do exist to address these ethical considerations?*

Our article is structured as follows: First, we briefly introduce the utilized research approach. Second we provide a thorough overview of the analog roots of digital nudging including existing ethical considerations in the BE and psychological domain. This provides IS scholars with a more thorough and focused background. Third, we introduce the IS perspective and indicate that several mechanisms utilized in nudging have been utilized for different IS or human-computer interactions as well. Fourth, we

argue that what makes a mechanism a nudge is primarily defined by the ethical standards requested when using them. Lastly, we summarize our key findings and offer practical implications as well as future research directions.

2 Interdisciplinary Bridge-Building Approach

At the beginning of our research, we identified two important research considerations: (1) there is an extended body of knowledge regarding ethical nudging in BE research and (2) to the best of our knowledge no in-depth analysis of ethical digital nudging is present, taking the specialties of IS and digital environments into account. Referring to Cooper (1988), our research goal can be best described as integrating and transferring the research of one domain to another – eventually leading to an *interdisciplinary bridge-building* similar to Cooper's linguistic bridge-building: "bridging the gap between theories or disciplines" (Cooper, 1988). We take established concepts from a research domain/community A (in this case BE) as a starting point to discuss them in research domain/community B (IS / digital contexts). We argue that an overall perspective is an efficient, insightful and helpful starting point for this endeavor. As further research, scholars are then able to map more granular research with the overall context. As a result, we decided to adapt techniques from literature reviewing: We took central and pivotal ethical studies from domain A, summarized the key considerations and reviewed these against research domain B.

In sum, our research can be positioned to (1) focus on theoretical outcomes of previous research mainly in the BE domain, (2) integrate and synthesize the main thoughts of the BE domain to support linguistic bridge-building towards the IS domain, (3) provide an espoused position suggesting a synthesized understanding, (4) focus on central and pivotal articles, (5) be organized around concepts and similar abstract ideas, and (6) target the audience of specialized scholars (cf. Cooper, 1988). Selection-wise, we only considered peer-reviewed sources for our main analysis (deviating in specific circumstances, if for example seminal literature was published (e.g. Thaler and Sunstein (2008)) and if there is sufficient evidence that the authors are domain experts).

3 The Theoretical Foundation of Nudging

3.1 Human Behavior in the Mirror of (Behavioral) Economics and Psychology

For more than a century, scientists have been analyzing the determinants, patterns, and peculiarities in human decision making. Thereby, two fields of research stand out: social and cognitive psychology. The first mentioned strand analyzes how individuals' cognition, motivation, and behavior are influenced by societal factors, which include social preferences, self-schemes, social identity, and social norms. For instance, extensive empirical evidence illustrates that individuals do not act purely in self-interest, but follow social preferences like fairness, reciprocity, or an inequity aversion (Fehr and Gächter, 2002). Furthermore, their decisions are influenced by observed or implied behavior of other individuals or their peer-group (social norms) (Demarque et al., 2015).

The second strand of cognitive psychology focuses on internal thought processes. Building on the ground-breaking work of Kahneman and Tversky, the empirical research of the last decade documents numerous systematic patterns of human behavior, including heuristics and biases that can render decisions irrational (Kahneman et al., 1991; Kahneman and Tversky, 1979; Tversky and Kahneman, 1974; Gigerenzer and Todd, 1999; Loewenstein and Prelec, 1992).

Heuristics can be defined as simple information-processing rules, systematically used to substitute complex problems of decision-making by easier ones (Sanjari et al., 2017; Ross, 2014). Usually, heuristics refer to automatic, intuitive judgments, but it can also be utilized with deliberation, particularly when dealing with limited information. In many circumstances, such mental shortcuts preserve cognitive resources and enable faster and more efficient decision-making. Therefore, the use of heuristics is not irrational at all under various circumstances, especially in everyday decisions (Evans, 2008; Hutchinson and Gigerenzer, 2005). However, depending on the context, heuristics can also lead to unreliable

judgments and errors in decision-making. These predictable and systematic patterns of perceptual distortion, incorrect judgment, or illogical interpretation are called cognitive biases (Haselton et al., 2005; Tversky and Kahneman, 1974).

It is very common in psychology to explain such intuitive, automatic and affective modes of human decision-making via a dual process model (Lades, 2014; Kahneman, 2011; Stanovich and West, 2000). Accordingly, cognitive processes can be divided into two separate but interacting systems: An intuitive, automatic, emotion-driven, and quickly operating "System 1" on the one hand, and a reflective, logically calculating, but slower processing "System 2" on the other hand (Lades, 2014; Kahneman, 2011). Although the notion of "System 1/2" is widespread in BE, this does not suggest that "the two types of processes are located in just two specific cognitive or neurological systems" (Evans and Stanovich, 2013). Accordingly, they suggested to reframe "System 1/2" to "Type 1/2". The task, the specific context, and the person's cognitive capacities determine which system will be addressed in a choice decision. This process usually occurs unconsciously, which is an important prerequisite to the effectiveness of the nudging mechanism. Since both systems run parallel and frequently interact with each other, the boundaries between them maintain a fluency. Consequently, individuals respond systematically, in some instances affectively and subconsciously, to physical, verbal, emotional, or social stimuli and other "seemingly irrelevant factors" (Thaler, 2016), which include positioning effects, default rules, framing descriptions, and accounting (Thaler, 2016).

Likewise and contradictory to the economic paradigm of rational choice, numerous empirical observations have shown that several anomalies in human (rational) decision-making do not occur sporadically, but rather in a systematic and predictable manner (Sunstein, 2011). These empirical insights paint a more realistic picture of human behavior: Humans have only limited cognitive resources and their judgments and decisions are strongly influenced by emotions, social norms, the use of mental shortcuts (heuristics), biases, as well as through the context of the decision per se (Heukelom, 2014).

Nevertheless, it took decades to integrate these systematic and predictable deviations from rational behavior and henceforth into a more empirical and behaviorally informed policy approach (Reisch and Zhao, 2017; Sunstein, 2011; OECD, 2017; Lunn, 2014; Dolan et al., 2012). One substantial part of this process posed is the development of empirically informed, less intrusive, and choice-preserving behavioral change intervention mechanisms: termed as so-called *nudges* (Reisch and Zhao, 2017; Thaler and Sunstein, 2008, 2003).

3.2 Definition of Nudging

So, what constitutes a nudge? The definition of Sunstein and Thaler (2008) provides a starting point for the following discussion: "A nudge [...] is any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives" (Thaler and Sunstein, 2008). From a BE perspective, every decision-making process is embedded in some kind of "choice architecture". This omnipresent and exogenously defined informational or physical environment structure can influence the manner of how individuals behave and what decisions are made (Thaler and Sunstein, 2008; Lehner et al., 2016; Johnson et al., 2012).

Barton and Grüne-Yanoff (2015) point out that there is a dualistic nature inherent to this standard definition. Depending on the point of view, a nudge can either be interpreted as a non-intentional aspect of choice architecture (Mills, 2015), or as a deliberately and purposefully implemented intervention tool (Guala and Mittone, 2015). To resolve this controversy, Sunstein suggests to clearly distinguish between the basic idea of choice architecture on the one hand, and nudges as an intentional and targeted intervention to modify choice architectures on the other hand (Sunstein, 2015b). For a detailed overview of different choice architecture techniques, refer to Münscher et al. (2016).

However, even this differentiation is too vague for a highly selective and clear distinction since "almost everything that affects behavior has been renamed a nudge, which renders this concept meaningless" (Gigerenzer, 2015). Moreover, even the payoff function of a rational agent is not exclusively determined by economic incentives (Hansen and Jespersen, 2013). Consequently, this means the core of the definition needs to be refined (Barton and Grüne-Yanoff, 2015).

For this reason, Hausman and Welch (2010) suggest to broaden the definition to encompass all other types of incentives as well: "Nudges are ways of influencing choice without limiting the choice set or making alternatives appreciably more costly in terms of time, trouble, social sanctions, and so forth" (Hausman and Welch, 2010). However, a link to the underlying cognitive processes sustained as a key component in the development of nudge-interventions is still lacking.

To reconcile the definition and its theoretical underpinnings, Hansen (2016) has outlined an extended definition which highlights the importance of BE principles: "A nudge is a function of (I) [sic] any attempt at influencing people's judgment, choice, or behavior in a predictable way, that is (1) made possible because of cognitive boundaries, biases, routines, and habits in individual and social decision-making posing barriers for people to perform rationally in their own self-declared interests, and which (2) works by making use of those boundaries, biases, routines, and habits as integral parts of such attempts [...]" (Hansen, 2016).

The fact that nudge-interventions use such behavioral insights to influence peoples' behavior is a central part in the ethical criticism in regards to the approach in both academic and public sphere (Hansen, 2016; Felsen et al., 2013). A fundamental and widespread claim is that nudges would work solely by manipulating people's choices in a subconscious manner, incompatible with the basic concept of modern democracies (Vallgårda, 2012; Hansen and Jespersen, 2013; Bovens, 2009).

Indeed, behavioral interventions, triggering automated and affective-cognitive processes (e.g. heuristics) in order to alter people's behaviors directly, resemble the core of the nudging-approach (Hansen and Jespersen, 2013; Barton and Grüne-Yanoff, 2015). These interventions are designed to unconsciously influence people's behavior by exploiting biases (biasing) or using a second bias to compensate the effects of another bias (re-biasing) (Soman and Liu, 2011; Michalek et al., 2016). But as Michalek et al. (2016) underline, nudges do not necessarily address these subconscious judgement processes (nondeliberative faculties). Instead, nudges can also evoke behavior change in an indirect way by stimulating reflective decision-making processes. These tools consciously "de-bias" people's judgments by correcting or eliminating cognitive biases or "monitoring errors" of system 2 (Michalek et al., 2016; Barton and Grüne-Yanoff, 2015; Hansen and Jespersen, 2013).

Regardless of an individual's evaluation to this controversy, the discourse emphasizes that there have always been political, ethical, and cultural factors inherent to and influencing nudge-interventions (Hagman et al., 2015). Thaler and Sunstein reply to such criticisms by formulating some ethical principles for the usage of nudges. Accordingly, "to count as a mere nudge, an intervention has to preserve freedom of choice" (Thaler and Sunstein, 2008), must be transparent (Sunstein, 2015c), and "influence choices in a way that will make choosers better off, as judged by themselves" (Thaler and Sunstein, 2008). Following these especially ethical boundaries, intervention mechanisms can therefore become nudges, with these central ethical considerations discussed in the following chapter.

4 Ethical Considerations of Nudging

In this chapter, we derive 5 questions regarding important ethical considerations of nudging. These serve as the foundation for the ethical discussion considering digital contexts and IS following in chapter 6.

4.1 Autonomy / Freedom of Choice

The first ethical premise is that each nudge-intervention must be easy to avoid (e.g. by a simple mouse click). To qualify as a nudge, an intervention mechanism needs to preserve the individual's freedom of choice, interpreted as the absence of obstacles, barriers or constraints (Thaler and Sunstein, 2008, 2003): "A person has freedom of choice if he lacks constraints on the reasoned selection and performance of one more of the items on an action-menu" (Carter, 2004). According to such an "opportunity concept" (Carter, 2004), real freedom of choice can only exist if individuals are capable of making "reasonable" decisions and – at the same time – their choice options are not restricted by third parties.

However, it is controversial to whatever extent this criterion becomes sufficient to protect the individual's autonomy and freedom of choice: "Yet those who have been worried about the ways in which government action and social pressure limit liberty have been concerned about liberty in a wider sense than closing off alternatives or rendering them costlier. Let us call the other aspects of this wider sense of liberty, 'autonomy' – the control an individual has over his or her own evaluations and choices." (Veetil, 2011). Considering this aspect, it becomes clear that in order to highlight the individual's freedom of choice, it is not a sufficient condition to esteem the (process) value of freedom. The crucial point is even "if [a] choice architecture does not block or significantly burden choices, it might still interfere with a person´s ability to discern and consider options and act according to her own preferences, i.e., it might interfere with her autonomy" (Blumenthal-Barby, 2013).

According to Sunstein (2014c), this criticism neglects that in modern societies it is very difficult for individuals to make autonomous and well-founded decisions due to the reason they are permanently exposed to numerous different environmental stimuli. The effort necessary to discern a nudge needs to be compatible with mental (computational) capacities of individuals (Simon, 1955). Depending on how individuals perceive the situation or system in which the nudge occurs, the required effort to recognizes a nudge may vary. Nudge-interventions should not be equated with a limitation of autonomy, as they have the ability to offer guidance, especially in complex social or economic contexts. In this manner, nudges provide an opportunity to improve the ability to make informed and autonomous decisions. In addition, nudges can prevent errors in judgments by stimulating reflective decision-making processes, e.g. utilizing feedback mechanisms (Sunstein, 2015c).

Nevertheless, it shall be noted that even the act of ignoring an otherwise entirely ethical nudge, necessitates cognitive capacities (which are limited), thus incurring psychological costs ("psychic tax") for nudgees while they sustain their freedom of choice (Barton and Grüne-Yanoff, 2015; Bovens, 2009). This leads to the following ethical question:

> *E1: How much effort on the behalf of individuals is justified to preserve their freedom of choice?*

4.2 Transparency

A further opposition manifested towards nudging is that nudges may impose a "lack of transparency by clouding the availability of options accessible at the moment of the decision and thereby endanger informed and free choice" (Clavien, 2018). Hence it is important to save and preserve nudgees' freedom of choice, nudgees must be able to easily recognize when and where they are subject to being nudged (type-transparency), as well as what the nudger's goals of this intervention are, in addition to how and why the nudge is working (token-transparency) (Bovens, 2009; Barton and Grüne-Yanoff, 2015). Accordingly, the ethical question is:

> *E2.1: How much concealment of a nudge is bearable to still be considered transparent?*

As an approach to this consideration, it is argued in favor of certain public regulations that ensure a legitimate usage of nudging interventions: "We should create rules of engagement that reduce fraud and other abuses, that promote healthy competition, that restrict interest-group power, and that create incentives to make it more likely that [nudging, the authors] architects will serve the public interest" (Thaler and Sunstein, 2008). Following the transparency rule, nudges have to be clearly identifiable by nudgees (Sunstein, 2015c). As a starting point in order to classify nudges as transparent, the schemata of Hansen and Jespersen (2013) may be referred to, while extended criteria to e.g. measure the transparency of a nudge are still missing. The ethical question is:

> *E2.2: How much difficulty to identify nudges as nudges is justified in order to still consider the nudge as transparent?*

4.3 Goal-Oriented Justification

Nudges can be targeted at three, not necessarily exclusive, goals (Clavien, 2018; Hagman et al., 2015): selfish goals of the nudger (e.g. own benefit, profit), pro-social / social goals (e.g. public welfare, gender

equality) and pro-self / nudgee-driven goals (e.g. participating more in sports if the individual has already formed an attitude for this). However, according to Thaler and Sunstein (2008), there is no ethical justification especially for the first goal (selfish), as the selfish perspective does not necessarily "influence people's behaviour in order to make their lives longer, healthier, and better" (ib.).

Regarding the pro-self perspective, various authors have pointed out that the practical implementation of this legitimation strategy suffers from a lack of knowledge: It seems presumably impossible for the nudger to identify the preferences of the individual (Barton and Grüne-Yanoff, 2015; Rebonato, 2014). Moreover, even if one ignores this point of criticism, the challenge still remains to create one single intervention that fits all heterogeneous preferences within a large population. Because of this heterogeneity, there are always going to be nudges that steer some individuals away from their own true preferences (Barton and Grüne-Yanoff, 2015). Due to this criticism, the pro-self nudge has to stem from justifiable and ethically acceptable goals (Ismaili M'hamdi et al., 2017), enabling nudgees to avoid "decisions they would not have made if they had paid full attention and possessed complete information, unlimited cognitive abilities, and complete self-control" (Hagman et al., 2015). This leads to the ethical question:

> *E3.1: How aligned do choice architects' goals need to be with those of the individuals' in order to render a nudge as justifiable?*

Clavien (2018) determined four strands of argumentations typically used by practitioners to claim their nudges as ethically acceptable: (1) The nudge helps to achieve some desirable outcome or (2) to fulfil important values or moral principles. (3) The nudgers' goals stem from good intentions and (4) one can provide evidence that nudgees do or would share the goals pursued by the choice architect. From a moral point of view, the last argument is the most ethically sound as there is evidence provided (e.g. by representative surveys or interviews) that there is sufficient belief that nudgees do share the pursued goals of the choice designer. Although the first three argumentations may be benevolently or hedonistically motivated, the conflict remains whether the choice architects' goals align with the goals of the nudgees and whether their autonomy is preserved.

Regarding the pro-social perspective, nudges can even be justified if individuals have not actively consented (for example through governmental programs to increase individuals' health or retirement savings). However, "nudges will leave a trace of moral violation as long as there are reasons to think that some nudgees would find it unbearable to be nudged" (Clavien, 2018). In similar vein, some nudgees may disagree with the nudgers' goals or disagree with being nudged at all. With a complete consensus of all potential nudgees remaining mostly difficult to attain, the ethical question arises:

> *E3.2: How much disagreement among targeted individuals is bearable to still justify a nudge's pro-social implementation?*

Clavien (2018) identified four ethical replies to this question: (1) Substantiate the importance of the choice designer's goals, (2) downplaying the ethical concerns of nudging with respect to more negative or unregulated means (e.g. advertisements), (3) arguing that the nudgees' goals are in line or shared by the choice architect or (4) stating that nudgees do or would agree with being nudged (either by consent or if being asked). Defenders of a libertarian paternalistic view usually take the third perspective arguing in favour of allowing nudgees to make "better" decisions (Thaler and Sunstein, 2008).

5 The Advent of Digital Nudging

5.1 Nudging in Digital Choice Environments

The key to understanding, developing and implementing nudges lies in the foundational psychological determinants of human decision-making. First empirical results have illustrated that digital choice environments stimulate a rather automatic and intuitive mode of thinking (system 1), due to their higher visuality and the pure mass of available information (Benartzi and Lehrer, 2015). Furthermore, research has indicated that information-abundant digital environments lead to choice overloads and decreasing sustained attention spans, thus individuals are spending less time with in-depth and concentrated reading

on digital screens (Liu, 2005). Neuroscientists spotted that especially younger people tend to "gravitate toward "shallow" [digital, the authors] information processing behaviors characterized by rapid attention shifting and reduced deliberation. They engage in increased multitasking behaviors that are linked to increased distractibility and poor executive control abilities" (Loh and Kanai, 2016).

Under these attention-limiting circumstances, individuals are more vulnerable to judgment errors when processing digital information. This concerns both "traditional biases" (e.g. priming or anchoring effects) as well as digital-specific phenomena, like the so-called "display-biases" (Benartzi and Lehrer, 2015). Consequently, the pure arrangement of options (Christenfeld, 1995) or their visual salience (Milosavljevic et al., 2012) can affect decision-making (Benartzi and Lehrer, 2015). For example, individuals are more likely to notice options or information if they are displayed in the middle of the screen ("middle bias"). Although this phenomenon also occurs in the "analog world", it is particularly pronounced in digital space due to the high visuality of screens (Benartzi and Lehrer, 2015; Reutskaja et al., 2011; Simons et al., 2017).

Besides these challenges, digital applications offer numerous innovative tools and functions with the potential to increase the effectiveness of nudging interventions. Among other things, these include filter options, recommender systems, intelligent tracking and targeting methods or feedback tools and a wide range of techniques for individualization and personalization (Mirsch et al., 2017). Compared to the analog and physical world, digital environments offer a higher degree of freedom for the digital nudge designer. While the physical world is present "as it is" and one cannot easily reshape or move trees, streets or the like, developers and designers usually own a variety of options to change the appearance of their IS without significant monetary efforts. Furthermore, the opportunity to change something from the current state into a preferred something (Simon, 1981) is easier to realise within digital environments (e.g. through digital prototyping). Therefore, implementing and testing digital nudges can be considered easier and faster than analog nudges.

5.2   The Concept of Digital Nudging

Recently within the IS domain, nudging has gained more attention. So far, digital nudging is conceptualized as "the use of user-interface design elements to guide people's behavior in digital choice environments" (Weinmann et al., 2016). Digital choice environments are user interfaces "that require people to make judgments or decisions" (ib.). This broad definition leaves some room for interpretation: What are the specific characteristics of a (digital) nudge and how does this approach relate to other research strands in IS, like persuasive technology (PT) or behavior change support systems (BCSS)?

The alteration of user behaviours also persists as part of PT and BCSS research. As is the case with digital nudging, the notion of these concepts likewise is rather broad: PT as "any interactive computer system designed to change people's attitudes or behaviors" (Fogg, 2003) and BCSS as IS "designed to form, alter or reinforce attitudes, behaviors or an act of complying without using coercion or deception" (Oinas-Kukkonen, 2013). Following these broad definitions for digital nudging, PT and BCSS, there seems to be no clear demarcation. Accordingly, nudging is seen as a part of both PT and BCSS (Schneider et al., 2018). However, PT and BCSS intervention mechanisms are not automatically a part of digital nudging as the prevailing and vague definition of digital nudging based on Weinmann et al. (2016) may suggest.

In line with Meske and Potthoff (2017), we argue that a general digital intervention mechanism becomes a nudge by fulfilling the definitory and especially ethical strings attached to a nudging intervention mechanism rooted in the BE domain. By adhering to both definitory and ethical standards introduced in chapter 3 and 4, we argue for digital nudging to be the digital representation of nudges as understood in BE. Thus, only some intervention mechanisms of PT or BCSS qualify as nudges.

This approach (rooting digital nudging in BE research) allows IS scholars to profit from the vast research available in BE (and partly psychology) on nudging. Scholars can sharpen the concept of digital nudging and establish more research on the influence of IS on nudging, thus reaching at a more substantial un-

derstanding of digital nudging. In line with our argument to root digital nudging more in its BE provenance, we apply the ethical considerations on nudging to the specialties of digital environments in the following chapter.

6 Ethical Considerations of Digital Nudging

In this chapter, digital nudging is reasoned with respect to our guiding research questions stated in the introduction as well as the ethical considerations raised in chapter 4. Table 1 summarizes these ethical guiding questions. This reasoning is two-fold: First, we transfer the results from chapter 4 in order to address *RQ1: Which ethical considerations arise when using nudging mechanisms in digital choice environments (i.e. digital nudging)?* If necessary, adaptations or extensions of the ethical concerns are suggested. Second, potential approaches to address these ethical concerns amidst a digital context are introduced, answering *RQ2: What potential resolutions do exist to address these ethical considerations?*

Freedom of Choice / Autonomy
▪ E1: How much effort on the behalf of individuals is justified to preserve their freedom of choice?
Transparency
▪ E2.1: How much concealment of a nudge is bearable to still be considered transparent?
▪ E2.2: How much difficulty to identify nudges as nudges is justified in order to still consider the nudge as transparent?
Goal-Oriented Justification
▪ E3.1: How aligned do choice architects' goals need to be with those of the individuals' in order to render a nudge as justifiable?
▪ E3.2: How much disagreement among targeted individuals is bearable to still justify a nudge's pro-social implementation?

Table 1.   Guiding Questions Regarding Important Ethical Considerations for (Digital) Nudging

6.1 Freedom of Choice / Autonomy

> *E1: How much effort on the behalf of individuals is justified to preserve their freedom of choice?*

Digital nudges have to preserve the entire freedom of choice of individuals. To enable this, the autonomy of nudgees has to be preserved when being subject to digital nudging. The intervention must not prohibit or limit specific choices (e.g. offering a sports activity guideline on a website only if a user enters his e-mail-address is not a nudge, even if the intention of the document is to help users to do more sport activities). As research has pointed out, digital contexts are frequently perceived as information overloaded (Liu, 2005), leading to phenomena like shorter attention spans, banner blindness and shallow information processing behaviors (Loh and Kanai, 2016; Benway, 1998; Burke et al., 2005; Owens et al., 2011). This indicates that users are even less aware when being nudged in a digital context in contrast to analog choice environments. The choice architect thus has an even larger responsibility to not exploit this limited attention and choice overload in digital environments.

Another noticeable aspect is that digital nudges can be implemented more easily and at nearly no cost (Schneider et al., 2018) and choice architects may thus be tempted to excessively use these mechanisms. Even more, digital choice designers have higher degrees of freedom as they can theoretically modify every single element of a digital artifact like a website or app. For example, digital nudges can be used to remind individuals to demonstrate a certain behavior like doing more sports. These notifications (e.g. through an app) do not incur any implementation costs (the data costs are negligible) and as a result they may be at risk to over-nudge individuals (e.g. by sending hourly notifications). Coupled with this reason, not only the frequency but also the timing and the form of nudging can impact individuals' freedom of choice and overall autonomy (e.g. vibration or audio notifications cannot only disrupt meetings but also prevent individuals from fully concentrating themselves for a longer period of time). As a consequence, inappropriate timing and frequency may incur significant psychological costs ("psychic tax") through more cognitive capacities employed in dealing with these interruptions (Barton and Grüne-Yanoff,

2015). Prospectively, this may even support emerging phenomena such as smartphone addiction (Duke and Montag, 2017).

As a resolution, digital choice designer shall be especially sensitized towards their responsibility of designing nudge-interventions in a non-exploitative, non-prohibiting and choice-preserving way. New digital technologies shall be used, to preserve individuals' freedom of choice to the highest possible degree. For this maximum autonomy, individuals should be able to influence the type, form and frequency of the nudge-interventions they are faced with. For example, users could be able to configure their IS experience by choosing which kind of mechanisms they want to be subject to. In a web shop, this would mean that users can deactivate or personalize several functions like recommender systems (e.g. what have other customers bought, additional items presented in the shopping cart), reminder mechanisms (e.g. push notifications or website lay-overs), conversational agents (e.g. product specification chatbots), scarcity mechanisms (e.g. only X products left), simplification mechanisms (e.g. one-click-buying options that make ordering easier but also more tempting) or website personalization (e.g. personally greeting an user). Individuals shall be asked upfront about their favorite personal default options, that can be changed at any time. Furthermore, users shall be actively informed about updates to the IS, what these updates mean and whether novel default options are introduced, or existing ones removed.

In sum, as a certain amount of cognitive effort to neglect any kind of nudging mechanisms cannot be avoided, this cognitive effort shall be as insignificant as possible. Ideally, it is only a one-time effort to ignore a nudge (as it is not presented again). If the nudge is presented again and again, there must be a strong pro-self or pro-social argumentation present to justify itself as ethical. The necessary effort level to dismiss a digital nudge can be measured alongside with the transparency introduced in the following section.

6.2 Transparency

E2.1: How much concealment of a nudge is bearable to still be considered transparent?

While for analog nudges the token-transparency (regarding the transparency of the nudger's goals as well as how and why a nudge is working) might be explicated with a reasonable effort, full token-transparency becomes a major challenge in present IS. This is mostly due to the increasing complexity of IS and advancement of artificial intelligence algorithms. As the analog nudge, for example in a cafeteria to choose fruits instead of snacks by switching the positioning of both offerings is rapidly explained, digital nudges can be both easy to understand (e.g. if they follow an explicit and limited ruleset). However, if nudges are initiated and steered by a highly sophisticated (machine learning) algorithms, the understanding of why and how a nudging-algorithm is actually nudging becomes more and more a black box even to computer scientists. Machine learning algorithms have been established in classifying huge amount of data sets into different categories. Based on the classification, they mostly provide different outputs. In this study's context the decision may be to use a certain nudge or whether to nudge at all. This algorithm as "nudger" can become "opaque in the sense that if one is a recipient of the output of the algorithm (the classification decision), rarely does one have any concrete sense of how or why a particular classification has been arrived at from inputs" (Burrell, 2016).

As a resolution, we argue to carefully choose the design of a "digital choice architect". These considerations shall include ethical considerations of algorithms as well as ensuring the interpretability of the (machine learning) model and its outcomes (Mittelstadt et al., 2016; Vellido et al., 2012). To ensure token-transparency for nudgees, we see two approaches: (1) to limit the algorithmic complexity of the nudging procedures to a level that a typical individual can comprehend, or (2) to ensure the ethical justification through professional means, such as external auditors assessing the code as nudging-compliant (Pasquale, 2015) or to team up in interdisciplinary partnerships to ensure the ethical compliance of one's algorithmic endeavors (Burrell, 2016; Kraemer et al., 2011).

E2.2: How much difficulty to identify nudges as nudges is justified in order to still consider the nudge transparent?

Digital nudges have to be made transparent to nudgees as well. Full type-transparency in a digital context would mean that nudgees are able to easily recognize when and where they are subject to being nudged.

This raises several issues: As an easy recognition refers to the (visual) representation of a nudge, the question is how easy is easy enough. For example, is a border around a nudge or a hint, similar to the cookie-agreement confirmation, sufficient? Is an overall IS-wide information regarding the usage of nudges sufficient? And may nudges themselves be depicted as other interaction mechanisms (like buttons) or will they incorporate specific properties to become prominent to nudgees? Furthermore, is every non-monetary and non-forceful button or notification on a web-page a nudge?

As a resolution, we argue that intervention mechanisms qualifying as nudges need to exceed a certain threshold of goal setting and (human) consideration. Thus, a nudge is directed at something more than simply switching to the next webpage or interacting with a website normally. A conscious choice architect's goal setting has to be present in order to qualify interventions via any kind of user interface elements as nudges.

Finally, nudgees shall be able to identify the nudging intervention with a basic level of attention. Forms of highlighting a nudge may include borders, textual hints or other, easily recognizable means. Through online-based research (e.g. surveys), the noticeability of a nudging-design can be evaluated and confirmed already ex ante of implementing a nudge-intervention. Potential measures might be the time it takes nudgees to realize that they are nudged – or the probability that nudges are identified as such in a (simulated) stressful and choice-overloaded situation.

6.3 Goal-Oriented Justification

> *E3.1: How aligned do choice architects' goals need to be with those of the individuals' in order to render a nudge as justifiable?*

IS and digital environments share the same challenge as traditional BE contexts do – how to know what individuals want and how create one-size-fits-all interventions suitable to a variety of potentially targeted nudgees. Research has indicated that a general acceptance for nudges is present, but the acceptance differs with respect to the chosen intervention mechanism (Reisch and Sunstein, 2016; Sunstein, 2015a). However, digital environments provide additional tools as intervention mechanisms in order to stimulate behavior changes (Schneider et al., 2018). This environment may easily become more intrusive than analog environments as they introduce omnipresent and potentially more persuasive means to users (e.g. notifications or endless and always available information feeds): Choice architects are at a higher risk of being tempted in following their own (selfish goals) yet hiding their true ambitions behind a pro-social or pro-self façade. This eases even highly-unethically and manipulative means like "dark UX patterns", which are consciously misleading users in order to fulfil someone else's goals. Due to this and because IS designer can use less-expensive means to create persuasive systems while receiving more creative freedom in shaping interventions than physical designer, the goal setting of digital nudges is important to consider.

As a resolution, IS can provide three approaches: (1) less-expensive and broader user research tools, (2) more precise targeting mechanisms and (3) easier feedback mechanisms for individuals: (1) Regarding the user research tools, interviews and surveys have become a lot easier to implement through IS. Choice architects can hereby aim at a "shared preference justification" by arguing that they "pursue goals that are shared by, or in line with the preferences of individuals impacted by the nudge" (Clavien, 2018). Search engines and databases do also provide an easier access to available user research, allowing nudgers to understand their aspired nudgees in more depth. This precise understanding facilitates the ethical consideration process, as the pro-self (or pro-social) dimension can be easier justified if the target group is better understood and the alignment of goals from choice designers and nudgees is explicated. (2) Building on a solid understanding of their nudgees, choice architects can address the problem of one-size-fits all by specifying for which audience their nudge is justified: User targeting, differentiation and personalization offer more adjustment possibilities in digital than in analog contexts (Taken Smith, 2012; Kannan, 2017). For example, nudging people in a park to throw their rubbish into the designated bins via green footprints cannot be disclosed to certain park visitors. A digital nudge, however, may allow this differentiation (e.g. by displaying a digital banner only to a certain group of nudgees). This adaptation of digital marketing mechanisms in favor of supporting the design and implementation of

digital nudges can be an interesting field for future research. (3) Digital user research methods can also be integrated into IS to gather feedback from nudgees before, during or after being digitally nudged. These insights may render helpful in optimizing a nudging intervention and aligning it with the goals of nudgees.

> *E3.2: How much disagreement among targeted individuals is bearable to still justify a nudge's pro-social implementation?*

As mentioned before, it is often not feasible to reach an entire consent of nudgees. In these cases, certain nudges may still qualify as ethically acceptable if they fulfil additional ethical demands: Namely that nudges do not leave any trace of moral violation on nudgees (Clavien, 2018). Sunstein (2015c), for example, argues that social welfare nudges are mostly acceptable if choice architects can clearly argue why these are implemented and if all other specifics of a nudge are adhered to. In digital environments, the pro-social dimension can also be applied to legitimate nudging interventions. However, the conflict of a perfect consent among nudgees (mostly unfeasible) on the one hand and the pro-social argumentation to still legitimate interventions on the other hand remains.

As a resolution, we argue that digital user research and feedback tools shall be used to collect data on the consent of nudgees. Additionally, social and societal preferences can be recorded through these means, easing the understanding of the aspired target group of the intervention. This shall be part of pre-surveys and pre-tests such that the nudging intervention is already pre-evaluated when being introduced to a public audience. Resembling the previous section, choice designer can use digital environments to address a sub-target-group in order to increase the consent for a specific digital nudge among them. If no targeting is directly possible, nudgers may create a digital platform or network that users can voluntarily enter if they share the platforms goal (e.g. a platform that helps users to do sports more regularly). In this context, nudgers can obtain voluntary consent for a nudge by freely and transparently offering it.

7 Conclusion and Indications for Future Research

With the proliferation of digital choice environments, important ethical questions arise for an acceptable usage of digital nudges. We are approaching this challenge by bridging two domains: The nudging research stream in BE and the upcoming digital nudging stream in information systems research (ISR). Nudging-interventions have proven to be valuable tools in supporting individuals' behavior changes, but there is a price to pay for these: Interfering with the autonomy of individuals.

Thus, the supreme discipline would be to design (digital) nudges in a way that nudgees are openly willing to being nudged and all nudging specifics and ethical considerations are preserved.

These requirements prove an impossible endeavor as one may always find an individual being against a specific nudge or against being nudged overall (Clavien, 2018). Accordingly, scholars and practitioners have to understand when, why and how they may justify an ethical deviation from the "perfect" nudge (Sunstein, 2015c). This leads to the domain of ethics in arguing why these interventions may be good even if they interfere with individuals' autonomy. As a reply to Meske and Potthoff (2017) and Weinmann et al. (2016), we take care of this ethical realm in the context of digital nudge-interventions.

First, we carried out an interdisciplinary and integrative analysis of ethical nudging. Second, we summarized ethical considerations for nudges that have been raised mainly in BE and policy related research. Third, we transferred these considerations to the IS domain and indicated first approaches to these ethical concerns. Three important areas for ethical considerations are (1) freedom of choice and autonomy, (2) transparency and (3) goal-oriented justification of nudging. We argue that these concerns remain valid for digital environments and become even more important given the higher degree of freedom for choice architects and a growing toolkit of digital (persuasive) intervention mechanisms.

Concerning the evaluation of nudging ethics, these ethical considerations can hardly be addressed discretely or binary as there is usually no right or wrong (or a clear threshold demarcating both). Nevertheless, this does not relieve IS scholars and practitioners from assuming their responsibility, acting compliantly and thoroughly thinking about the ethical consequences of their actions. We support a greater awareness of important ethical considerations, not only in the context of digital nudging. This article

adds to the ethical discourse in the IS community in sum, and to the according ethical discourse for digital interventions in particular.

Taking our results into account, we argue that choice architects shall include ethical considerations already into their conceptualization and design phases of digital nudges in order to avoid ethically unsound nudges right from the start (ex-ante evaluation). Regarding digital policies, regulations for digital nudging may counterbalance the increasing interventional power of (digital) choice architects (e.g. through machine learning). At the same time, given the clear opportunities of digital interventions for individual and public welfare, enough freedom shall be granted to explore and test novel digital nudge-interventions. We feel confident that the extended body of ethical research in BE can render helpful for IS scholars, but we are aware of the fact that this article can just be one of the first steps towards an ethics of digital nudges. Bridging concepts and theories between different research domains provides new insights, avoids redundant research, but necessitated us to concentrate on pivotal and central articles, theories and concepts.

We conclude with several avenues for future research: First, we argue that a more thorough definition of digital nudging is necessary. The present definition of Weinmann et al. (2016) lacks important ethical concerns and shall be refined taking the reasoning of Hansen (2016) and the aspect of pro-social justification (Sunstein, 2015c) into consideration. Furthermore, the definition of digital nudging does not account for the IS epistemological notion of communication and dialogues, and should accordingly be integrated. Second, future researchers may take the myriad of application-oriented research into account that would have gone beyond the scope of this introductory article. One option can be to update the intervention mechanisms from BE to and by intervention mechansims within digital environments. Second, additional light shall be shed on extendingung the ethics of digital nudging and gaining more insights into the strength, weaknesses, opportunities and risks of digital nudging. Third, the aspect of digital nudge transparency evokes a tremendous challenge with respect to algorithmic transparency (if algorithms become digital choice architects) (Burrell, 2016), which shall be examined in more detail. Fourth, especially the three identified core elements of ethical nudging shall be evaluated in a digital context, in order to understand when and under which circumstances digitals nudges (1) preserve nudgees' freedom of choice, (2) are easily and reliably identified as nudges and (3) do meet nudgees' goals or (4) the goal-setting is such strongly legitimated by social goals that an intervention in (1), (2) and (3) can be justified. Fifth, the practical application of ethical considerations within the conceptualization, design and implementation phase of a digital nduge shall be further examined. An integration of practical ethical guidelines into existing frameworks (Meske and Potthoff, 2017; Michie et al., 2011) for the creation of nudges may help both IS praacitioners and scholars.

8 Acknowledgments

References

Barton, A. and T. Grüne-Yanoff (2015). "From Libertarian Paternalism to Nudging - and Beyond" *Review of philosophy and psychology* 6 (3), 341–359.

Benartzi, S., J. Beshears, K. L. Milkman, C. R. Sunstein, R. H. Thaler, M. Shankar, W. Tucker-Ray, W. J. Congdon and S. Galing (2017). "Should Governments Invest More in Nudging?" *Psychological science* 28 (8), 1041–1055.

Benartzi, S. and J. Lehrer (2015). *The Smarter Screen. Surprising Ways to Influence and Improve Online Behavior.* New York: Portfolio/Penguin.

Benway, J. P. (1998). "Banner Blindness: The Irony of Attention Grabbing on the World Wide Web" *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 42 (5), 463–467.

Blumenthal-Barby, J. S. (2013). "Choice Architecture: A mechanism for improving decisions while preserving liberty?". In C. Coons and M. E. Weber (eds.) *Paternalism. Theory and practice*, pp. 178–196. Cambridge: Cambridge University Press.

Bovens, L. (2009). "The Ethics of Nudge". In T. Grüne-Yanoff and S. O. Hansson (eds.) *Preference Change. Approaches from Philosophy, Economics and Psychology*, pp. 207–219. Dordrecht, Heidelberg, London, New York, NY: Springer.

Burke, M., A. Hornof, E. Nilsen and N. Gorman (2005). "High-cost banner blindness" *ACM Transactions on Computer-Human Interaction* 12 (4), 423–445.

Burrell, J. (2016). "How the machine 'thinks': Understanding opacity in machine learning algorithms" *Big Data & Society* 3 (1), 205395171562251.

Carter, I. (2004). "Choice, freedom, and freedom of choice" *Social Choice and Welfare* 22 (1), 61–81.

Christenfeld, N. (1995). "Choices from Identical Options" *Psychological science* 6 (1), 50–55.

Clavien, C. (2018). "Ethics of nudges: A general framework with a focus on shared preference justifications" *Journal of Moral Education* 47 (3), 366–382.

Cooper, H. M. (1988). "Organizing Knowledge Syntheses: A Taxonomy of Literature Reviews" *Knowledge, Technology & Policy* 1 (1), 104–126.

Demarque, C., L. Charalambides, D. J. Hilton and L. Waroquier (2015). "Nudging Sustainable Consumption. The Use of Descriptive Norms to Promote a Minority Behavior in a Realistic Online Shopping Environment" *Journal of Environmental Psychology* 43, 166–174.

Dolan, P., M. Hallsworth, D. Halpern, D. King, R. Metcalfe and I. Vlaev (2012). "Influencing Behaviour: The Mindspace Way" *Journal of Economic Psychology* 33 (1), 264–277.

Duke, É. and C. Montag (2017). "Smartphone addiction, daily interruptions and self-reported productivity" *Addictive behaviors reports* 6, 90–95.

Evans, J. S. B. T. (2008). "Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition" *Annual review of psychology* 59, 255–278.

Evans, J. S. B. T. and K. E. Stanovich (2013). "Dual-Process Theories of Higher Cognition: Advancing the Debate" *Perspectives on psychological science : a journal of the Association for Psychological Science* 8 (3), 223–241.

Felsen, G., N. Castelo and P. B. Reiner (2013). "Decisional enhancement and autonomy: public attitudes towards overt and covert nudges" *Judgment & Decision Making* 8 (3).

Fogg, B. J. (2003). *Persuasive Technology. Using Computers to Change What We Think and Do.* San Francisco: Morgan Kaufmann.

Gigerenzer, G. (2015). "On the Supposed Evidence for Libertarian Paternalism" *Review of philosophy and psychology* 6 (3), 361–383.

Gigerenzer, G. and P. M. Todd (1999). *Simple Heuristics That Make Us Smart.* New York: Oxford University Press.

Goes, P. B. (2013). "Editor's Comments: Information Systems Research and Behavioral Economics" *MIS Quarterly* 37 (3), iii–viii.

Guala, F. and L. Mittone (2015). "A Political Justification of Nudging" *Review of philosophy and psychology* 6 (3), 385–395.

Hagman, W., D. Andersson, D. Västfjäll and G. Tinghög (2015). "Public Views on Policies Involving Nudges" *Review of philosophy and psychology* 6 (3), 439–453.

Hansen, P. G. (2016). "The Definition of Nudge and Libertarian Paternalism: Does the Hand Fit the Glove?" *European Journal of Risk Regulation* 7 (01), 155–174.

Hansen, P. G. and A. M. Jespersen (2013). "Nudge and the Manipulation of Choice" *European Journal of Risk Regulation* 4 (01), 3–28.

Haselton, M. G., D. Nettle and P. W. Andrews (2005). "The Evolution of Cognitive Bias". In D. M. Buss (ed.) *The Handbook of Evolutionary Psychology*, pp. 724–746. Hoboken, N.J.: John Wiley & Sons.

Hausman, D. M. and B. Welch (2010). "Debate: To Nudge or Not to Nudge" *Journal of Political Philosophy* 18 (1), 123–136.

Heukelom, F. (2014). *Behavioral Economics*. New York: Cambridge University Press.

Hummel, D., P. Toreini and A. Maedche (2018). "Improving Digital Nudging Using Attentive User Interfaces: Theory Development and Experiment Design" *DESRIST 2018 Proceedings (Forthcoming)*.

Hutchinson, J. M. C. and G. Gigerenzer (2005). "Simple Heuristics and Rules of Thumb: Where Psychologists and Behavioural Biologists Might Meet" *Behavioural processes* 69 (2), 97–124.

Ismaili M'hamdi, H., M. Hilhorst, E. A. P. Steegers and I. de Beaufort (2017). "Nudge me, help my baby: on other-regarding nudges" *Journal of medical ethics* 43 (10), 702–706.

Johnson, E. J., S. B. Shu, B. G. C. Dellaert, C. Fox, D. G. Goldstein, G. Häubl, R. P. Larrick, J. W. Payne, E. Peters, D. Schkade, B. Wansink and E. U. Weber (2012). "Beyond Nudges: Tools of a Choice Architecture" *Marketing Letters* 23 (2), 487–504.

Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Penguin Books.

Kahneman, D., J. L. Knetsch and R. H. Thaler (1991). "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias" *Journal of Economic Perspectives* 5 (1), 193–206.

Kahneman, D. and A. Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk" *Econometrica* 47 (2), 263.

Kannan, P. K. (2017). "Digital marketing: A framework, review and research agenda" *International Journal of Research in Marketing* 34 (1), 22–45.

Kraemer, F., K. van Overveld and M. Peterson (2011). "Is there an ethics of algorithms?" *Ethics and Information Technology* 13 (3), 251–260.

Lades, L. K. (2014). "Impulsive Consumption and Reflexive Thought: Nudging Ethical Consumer Behavior" *Journal of Economic Psychology* 41, 114–128.

Lehner, M., O. Mont and E. Heiskanen (2016). "Nudging – A promising tool for sustainable consumption behaviour?" *Journal of Cleaner Production* 134, 166–177.

Liu, Z. (2005). "Reading behavior in the digital environment" *Journal of Documentation* 61 (6), 700–712.

Loewenstein, G. and D. Prelec (1992). "Anomalies in Intertemporal Choice: Evidence and an Interpretation" *The Quarterly Journal of Economics* 107 (2), 573–597.

Loh, K. K. and R. Kanai (2016). "How Has the Internet Reshaped Human Cognition?" *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry* 22 (5), 506–520.

Lunn, P. (2014). *Regulatory Policy and Behavioural Economics.* Paris: OECD Publishing.

Meske, C. and T. Potthoff (2017). "The DINU-model–a process model for the design of nudges" *n Proceedings of the 25th European Conference on Information Systems (ECIS)* 25, 2587–2597.

Michalek, G., G. Meran, R. Schwarze and Y. Özgur (2016). "Nudging as a New "Soft" Tool in Environmental Policy. An Analysis Based on Insights from Cognitive and Social Psychology" *Zeitschrift für Umweltpolitik & Umweltrecht* 8 (4), 169–207.

Michie, S., M. M. van Stralen and R. West (2011). "The behaviour change wheel: a new method for characterising and designing behaviour change interventions" *Implementation science : IS* 6, 42.

Mills, C. (2015). "The Heteronomy of Choice Architecture" *Review of philosophy and psychology* 6 (3), 495–509.

Milosavljevic, M., V. Navalpakkam, C. Koch and A. Rangel (2012). "Relative Visual Saliency Differences Induce Sizable Bias in Consumer Choice" *Journal of Consumer Psychology* 22 (1), 67–74.

Mirsch, T., C. Lehrer and R. Jung (2017). "Digital Nudging. Altering User Behavior in Digital Enviroments". In: *Proceedings of the 13. Internationale Tagung Wirtschaftsinformatik*, pp. 634–648.

Mittelstadt, B. D., P. Allo, M. Taddeo, S. Wachter and L. Floridi (2016). "The ethics of algorithms: Mapping the debate" *Big Data & Society* 3 (2), 205395171667967.

Münscher, R., M. Vetter and T. Scheuerle (2016). "A Review and Taxonomy of Choice Architecture Techniques" *Journal of Behavioral Decision Making* 29 (5), 511–524.

OECD (2017). *Behavioural Insights and Public Policy*. *Lessons from Around the World.* Paris: OECD Publishing.

Oinas-Kukkonen, H. (2013). "A Foundation for the Study of Behavior Change Support Systems" *Personal and ubiquitous computing* 17 (6), 1223–1235.

Owens, J. W., B. S. Chaparro and E. M. Palmer (2011). "Text advertising blindness: the new banner blindness?" *Journal of usability studies* 6 (3), 172–197.

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information:* Harvard University Press.

Rebonato, R. (2014). "A Critical Assessment of Libertarian Paternalism" *Journal of Consumer Policy* 37 (3), 357–396.

Reisch, L. A. and M. Zhao (2017). "Behavioural Economics, Consumer Behaviour and Consumer Policy: State of the Art" *Behavioural Public Policy* 1 (02), 190–206.

Reisch, L. A. and C. R. Sunstein (2016). "Do Europeans Like Nudges?" *Judgment and Decision Making* 11 (4), 310–325.

Reutskaja, E., R. Nagel, C. F. Camerer and A. Rangel (2011). "Search Dynamics in Consumer Choice Under Time Pressure: An Eye-Tracking Study" *American Economic Review* 101 (2), 900–926.

Ross, H. J. (2014). *Everyday Bias*. *Identifying and Navigating Unconscious Judgments in Our Daily Lives.* Lanham: Rowman & Littlefield Publishers.

Sanjari, S. S., S. Jahn and Y. Boztug (2017). "Dual-Process Theory and Consumer Response to Front-of-Package Nutrition Label Formats" *Nutrition reviews* 75 (11), 871–882.

Schneider, C., M. Weinmann and J. Vom Brocke (2018). "Digital Nudging. Guiding Online User Choices through Interface Design" *Communications of the ACM* 61 (7), 67–73.

Simon, H. A. (1955). "A behavioral model of rational choice" *The Quarterly Journal of Economics* 69 (1), 99–118.

Simon, H. A. (1981). *The Sciences of the Artificial.* 2. ed. Cambridge, Massachusetts: MIT Press.

Simons, A., M. Weinmann, M. Tietz and J. Vom Brocke (2017). "Which reward should I choose? Preliminary evidence for the middle-option bias in reward-based crowdfunding". In: *Proceedings of the 50th Hawaii International Conference on System Sciences*.

Soman, D. and M. W. Liu (2011). "Debiasing or Rebiasing? Moderating the Illusion of Delayed Incentives" *Journal of Economic Psychology* 32 (3), 307–316.

Stanovich, K. E. and R. F. West (2000). "Individual Differences in Reasoning: Implications for the Rationality Debate?" *The Behavioral and brain sciences* 23 (5), 645-726.

Sunstein, C. R. (2011). "Empirically Informed Regulation" *University of Chicago Law Review* 78 (4), 1349–1430.

Sunstein, C. R. (2015a). "Do People Like Nudges?" *SSRN Electronic Journal*.

Sunstein, C. R. (2015b). "Nudges, Agency, and Abstraction: A Reply to Critics" *Review of philosophy and psychology* 6 (3), 511–529.

Sunstein, C. R. (2015c). "The Ethics of Nudging" *Yale Journal on Regulation* 32.

Szaszi, B., A. Palinkas, B. Palfi, A. Szollosi and B. Aczel (2018). "A Systematic Scoping Review of the Choice Architecture Movement: Toward Understanding When and Why Nudges Work" *Journal of Behavioral Decision Making* 31 (3), 355–366.

Taken Smith, K. (2012). "Longitudinal study of digital marketing strategies targeting Millennials" *Journal of Consumer Marketing* 29 (2), 86–92.

Thaler, R. H. (2016). "Behavioral Economics: Past, Present, and Future" *American Economic Review* 106 (7), 1577–1600.

Thaler, R. H. and C. R. Sunstein (2003). "Libertarian Paternalism" *American Economic Review* 93 (2), 175–179.

Thaler, R. H. and C. R. Sunstein (2008). *Nudge: Improving Decisions About Health, Wealth and Happiness.* Illustrated ed. New Haven: Yale University Press.

Thaler, R. H., C. R. Sunstein and J. P. Balz (2013). "Choice Architecture". In E. Shafir (ed.) *The Behavioral Foundations of Public Policy*, pp. 428–439. Princeton: Princeton Univ. Press.

Tversky, A. and D. Kahneman (1974). "Judgment under Uncertainty: Heuristics and Biases" *Science* 185 (4157), 1124–1131.

Vallgårda, S. (2012). "Nudge: A New and Better Way to Improve Health?" *Health Policy* 104 (2), 200–203.

Veetil, V. P. (2011). "Libertarian Paternalism is an Oxymoron: An Essay in Defence of Liberty" *European Journal of Law and Economics* 31 (3), 321–334.

Vellido, A., J. D. Martín-Guerrero and P. J. G. Lisboa (eds.) (2012). *Making machine learning models interpretable:* Citeseer.

Weinmann, M., C. Schneider and J. Vom Brocke (2016). "Digital Nudging" *Business & Information Systems Engineering* 58 (6), 433–436.

White, M. D. (2013). *The Manipulation of Choice. Ethics and Libertarian Paternalism.* Online-Ausg. New York: Palgrave Macmillan.