DESIGNING DIGITAL CHOICE ARCHITECTURE

Studies on the behavioral and attitudinal impact of digital nudges

PhD dissertation submitted to the Faculty of Economics and Business

Institute of Information Management (IMI) and Institute of Management (IMN)

University of Neuchâtel, Switzerland

for the PhD degree in Information Systems

by

Kristoffer Bergram

Approved by the dissertation committee:

Prof. Adrian Holzer, University of Neuchâtel (UNINE), co-supervisor

Prof. Valéry Bezençon, University of Neuchâtel (UNINE), co-supervisor

Dr. MER Paul Cotofrei, University of Neuchâtel (UNINE), head of committee

Prof. Johannes Schöning, University of St. Gallen (HSG), external expert

Prof. Dennis Lalanne, University of Fribourg (UNIFR), external expert

2024

Université de Neuchâtel
Faculté des sciences économiques

IMPRIMATUR POUR LA THÈSE

DESIGNING DIGITAL CHOICE ARCHITECTURE:

Studies on the behavioral and attitudinal impact of digital nudges

Kristoffer BERGRAM

UNIVERSITÉ DE NEUCHÂTEL

FACULTÉ DES SCIENCES ÉCONOMIQUES

La Faculté des sciences économiques,

sur le rapport des membres du jury

Prof. Adrian HOLZER, Université de Neuchâtel, co-directeur de thèse

Prof. Valéry BEZENÇON, Université de Neuchâtel, co-directeur de thèse

Dr. MER Paul COTOFREI, Université de Neuchâtel

Prof. Denis LALLANE, Département d'informatique, Université de Fribourg

Prof. Johannes SCHÖNING, Laboratoire HCI, Université de St. Gall

autorise l'impression de la présente thèse.

Neuchâtel, le 19 mars 2024

Le doyen

Peter Fiechter

Faculté des sciences économiques
Avenue du 1er-Mars 26
CH-2000 Neuchâtel
Email: secretariat.seco@unine.ch
Tel: +41 32 718 15 00
www.unine.ch/seco

DESIGNING DIGITAL CHOICE ARCHITECTURE

Studies on the behavioral and attitudinal impact of digital nudges

Preface

This dissertation is based on a set of articles presented as chapters. These chapters are preceded by an introduction to the research context and concluded by a summary of the research contribution that also echoes a set of paths for future research.

I have conducted the research in the dissertation from conception to its final presentation, apart from the fourth chapter where the lead contribution is equally shared with the first co-author. The research in this dissertation was performed under the supervision of Prof. Dr. Adrian Holzer and Prof. Dr. Valéry Bezençon. The chapters featuring user studies in the dissertation received research ethics approval from the University of Neuchâtel's Research Ethics Board.

Summary

This dissertation explores the design space of digital nudges and their effects on users in three application domains: online privacy, social media usage, and online charitable giving. After the introductory chapter, a systematic review of empirical studies on digital nudging is conducted, uncovering the current research landscape. Next, an online experiment and a descriptive field study is used to demonstrate the effectiveness of digital nudges for enhancing the online privacy awareness of users for popular online services. The next chapter uses an online field experiment to alter the digital choice architecture of a popular social media platform. This study evaluates restricted newsfeed diets and their impact on user behaviors and experiences. The fifth chapter is an online factorial experiment concerning charitable giving with real monetary stakes. This last study systematically investigates combinations of several digital nudges on users' online donation behaviors and their associated attitudes around manipulation and autonomy. The dissertation extends several contributions related to digital nudging and choice architecture. The systematic review illuminates several open research avenues for the domain at large that call for future research. The experiment related to online privacy awareness shows that digital nudges could elevate user interactions with terms and privacy policies, although without bolstering users' recollections of what they have agreed to. The field experiment concerning social media newsfeed diets demonstrates their impact regarding time spent on the social media platform, but these design interventions come with both positive and negative user experiences. The online factorial experiment related to charitable giving reveals that some combinations of digital nudges increase donations, but these particular nudges also invoke users' concerns about autonomy threats and an experienced sense of manipulation. Together, the chapters of the dissertation highlight the complex behavioral and attitudinal impact of digital nudges across several relevant application domains.

Keywords: *digital choice architecture, digital nudging, persuasive technology, user studies, human computer interaction, information systems*

Résumé

Cette thèse explore l'espace de conception des nudges numériques et leurs effets sur les utilisateurs dans trois domaines d'application : la vie privée en ligne, l'utilisation des médias sociaux et les dons de charité en ligne. Après le chapitre d'introduction, une revue systématique de la littérature liée aux études empiriques sur les nudges numériques est menée, révélant le paysage actuel de la recherche. Ensuite, une expérience en ligne et une étude descriptive sur le terrain sont utilisées pour démontrer l'efficacité des nudges numériques dans la sensibilisation à la protection de la vie privée en ligne pour les utilisateurs de certaines plateformes digitales. Le chapitre suivant utilise une expérience de terrain, menée en ligne, pour modifier l'architecture des choix numériques d'une plateforme de médias sociaux populaire. Cette étude évalue les régimes restreints des fils d'actualité et leur impact sur les comportements et les expériences des utilisateurs. Le cinquième chapitre est une expérience factorielle en ligne concernant les dons de charité avec des enjeux monétaires réels. Cette dernière étude examine systématiquement les combinaisons de plusieurs nudges numériques sur les comportements de don en ligne des utilisateurs et leurs attitudes associées en matière de manipulation et d'autonomie. La thèse présente plusieurs contributions liées aux nudges numériques et à l'architecture de choix. L'examen systématique met en lumière plusieurs pistes de recherche ouvertes pour le domaine qui appellent à des recherches futures. L'expérimentation liée à la confidentialité en ligne montre que les nudges numériques pourraient améliorer les interactions des utilisateurs avec les conditions et les politiques de confidentialité, sans toutefois renforcer leur mémorisation de ce qu'ils ont accepté. L'expérience sur le terrain concernant les régimes des fils d'actualité des médias sociaux démontre leur impact sur le temps passé sur la plateforme de médias sociaux, mais ces interventions de conception s'accompagnent d'expériences positives et négatives pour l'utilisateur. L'expérience factorielle en ligne relative aux dons de charité révèle que certaines combinaisons de nudges numériques augmentent les dons, mais ces nudges suscitent également des inquiétudes chez les utilisateurs quant aux menaces pesant sur leur autonomie et leur sentiment de manipulation. Ensemble, les chapitres de la thèse mettent en évidence l'impact complexe des nudges numériques sur le comportement et les attitudes dans plusieurs domaines d'application pertinents.

Mots-clés : *architecture du choix numérique, nudge numérique, technologie persuasive, études d'utilisateurs, interaction homme-machine, systèmes d'information*

Acknowledgements

Jeremy Collins writes that sometimes the most important journeys are not necessarily from east to west, or from ground to summit, but from heart to head. Between these two places we find our own voice[*]. A PhD is more akin to a journey from one head to another. If we continue this journey long enough, it is between all these different heads that we may find our dissertation.

A necessary but not sufficient element for this voyage is, of course: A great many heads. These heads belong to a set of people. Most, but probably not all of these people will now receive recognition. The reason for this statement comes back to the following limitation related to these acknowledgments. There will inevitably be some type II errors in this list of people, also known as false negatives. If I have forgotten some of you, dinner is on me.

I want to start by underlining that this dissertation benefited from not one, but two excellent supervisors: Adrian and Valéry. Thank you for giving me a shot in the first place. As professors, you really complement each other, and I sincerely thank you for your persistent openness, guidance, and engagement. I will miss my access to both of your minds.

I want to echo my gratitude to my co-authors: Aditya, Abdessalam, Manon, Marija, Louis, Paul, Tony, and again Adrian and Valéry. I also want to thank all of the reviewers whose incisive feedback undoubtedly improved this work.

A special thanks to the University of Neuchâtel at large. I have been lucky to cross paths with a long list of people across our two institutes. While some of the people below have come and gone, I thank you all for the good times we have shared over the years: Vladimir, Aditya, Alessio, Eliane, Paul, Selena, Pierluigi, Eric, Yao, Michael, Claudia, Niña, Bruno, Romain, Valérie, Catalin, Emma, Afshin, Martina, Philippe and Guiseppe.

[*] Jeremy Collins, Drawn: The Art of the Ascent [p. 176].

I also want to thank Sara and Eugenia for the administrative service you have provided for me and so many others at our institutes. Stephanie, thank you always being open to workshop my attitude over coffee as soon as times got tough. Mitch, thank you for all the great conversations, seeing you on a Thursday always made it feel more like a Friday. Ertuğrul, thank you for plowing a path of excellence at work while at the same time keeping the most chill demeanor ever. Manon, thank you for making this sometimes slow struggle a lot more genuine, humorous and enjoyable. Abdessalam, thank you for constantly going out of your way to be kind to me. You have been a good friend and a sharp collaborator throughout this period.

Finally, even though we have been separated by some mountains, a sea, and the occasional pandemic – I will not forget about some of my friends and family up north. Special thanks goes to Dan, Patrik, Klara and Cornel. Whether it be over long hikes, vertical wind tunnels, road trips, or just good wine, you have all picked me up in your own unique ways and helped put things into their proper perspective. To Ammi and Peter: Thank you for your honest reflections, kindness, and love. While I have missed you a lot over these years - you have at the same time always been there for me.

To Tiffany (who has arguably nudged me a lot more than I'd like to admit): Thank you for patiently and lovingly holding my hand, while allowing me to go my own way during this journey.

Contents

List of Abbreviations

BIS	Business Information System
DSRM	Design Science Research Methodology
FBM	Fogg's Behavioural Model
FOMO	Fear of Missing Out
HCI	Human-Cumputer Interaction
ICRC	International Committee of the Red Cross
Mturk	Amazon's Mechanical Turk
PDB	Privacy Dialog Box
PP	Privacy Policy
PSD	Persuasive Systems Design
PT	Persuasive Technology
ToS	Terms of Service
UI	User Interface
VUI	Voice User Interface

List of Tables

List of Figures

1 Introduction and Definition of the Design Space

The first misconception is that it is possible to avoid influencing people's choices.

– Richard H. Thaler & Cass R. Sunstein [1, p. 10]

Contents

1.1 Research Context

Designers of various stripes are trained to build a user interface (UI) with menus, buttons, feeds, sliders, input fields, and other design elements. In principle, a designer is tasked with building an interface that is intuitive, and that helps users complete desired tasks and goals in an efficient manner [2, 3]. Today, these tasks and goals may determine important aspects of our everyday lives. For example, around 30% of US adults report to have used or are currently using a dating app/site [4], and a majority of people in the EU are now conducting their banking affairs online[5]. Important choices such as who we find as a potential partner, or how we manage our finances are increasingly influenced by the work of various designers. When a designer works with any aspect of human–computer interaction (HCI), they are typically responsible for arranging the digital context in which their users make decisions [6]. In other words, they are designing a digital choice architecture for the user. Sometimes trivial details in this choice architecture can steer human choices and behaviors in unexpected directions. When a change in the choice architecture steers the user in a different direction, this is often called a *nudge* [1]. Before I delve deeper into digital nudges and choice architecture, a very brief background will be provided of the antecedent insights and ideas that led up to the current research context.

1.1.1 Towards nudging and choice architecture

The concepts of nudging and choice architecture evolved on the back of several decades of research into human cognition and decision-making. As with several other research fields, an important insight was possibly first articulated by Herbert Simon in the 1950s. He argued for a theoretical revision of the rational "homo economicus" [7]. Instead of assuming that human choices are governed by rational behavior, Simon argued for a theory with: "...a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms, including man, in the kinds of environments in which such organisms exist" [7, p. 99]. In other words, human decision-making had certain unrecognized constraints (some of which are innate to us, and some of which can be found in the decision environment we operate in). Simon later referred to this idea as "bounded rationality" [8, p. 196]. Notably, Tversky and Kahneman [9], among several others [10] empirically clarified what several of these constraints were, both in relation to the choice environment and to humans' cognitive processes. This empirical clarification laid the foundation for an influential research program in psychology known as the *heuristics and biases* approach [11]. One of the core tenants of this research program was that heuristics act as mental rules of thumb and biases act as weights for human decision-making. This meant that in certain situations humans would act in more or less predictably irrational ways [12].

To illustrate: human decision–making can be heavily influenced by social conformity or herd behavior. Even before various cognitive biases and heuristics had been thoroughly mapped by researchers, Asch [13] published the results of an experiment that stumbled upon what seemed to be a bug in human visual judgment. During each trial of Asch's experiment, a group of seven to nine students were shown a set of two cards, see Figure 1.1 on the next page. The students were asked to pick the line on the second card that has the same length as the line on the first card [13]. In reality, only one member of the group of students was a true subject – the other members were an instructed majority and were part of the experimental setup. According to a variable schedule, during two thirds of the repeated trials this instructed majority would erroneously pick the wrong line on the second card while being in the same room as the participant.

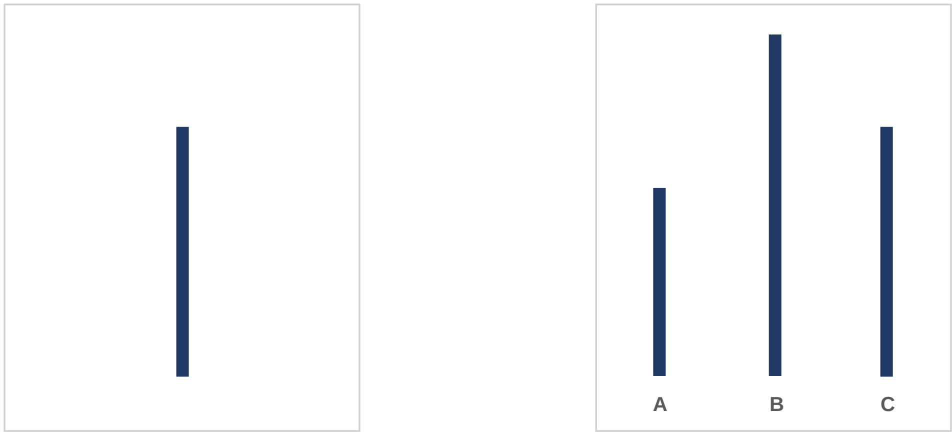

Figure 1.1: Examples of the two cards that participants were shown during Asch's experiments on social conformity. Modified and adopted from Mojzisch and Krug [14].

When participant's choices were completely separated from the choices of others, the error-rate when it came to picking the correct line was less than 1%, but when the erroneous judgments of the instructed majority were present, this error-rate reached nearly 37 % [13]. Based on Figure 1.1, it appears rather difficult to make an error given the task at hand, but the main result from this experiment still replicates for adults and children alike without using an instructed majority of actors [15, 16]. By instead relying on digital technology such as color–coded PowerPoint slides, researchers can re-create the impression of a minority vs. majority situation when people make their visual judgments. This is accomplished by randomly distributing polarized glasses among participants so that the length of the lines (like those in Figure 1.1) will change depending on what type of glasses the participants are wearing [15, 16]. The lesson from Asch's experiments is that the cues from the herd can have a strong influence on peoples' judgments. There are many other examples like the one above, but the main point is that over several decades researchers have shown how human behavior and choice can be unexpectedly influenced by a variety of factors. Insightful examples include: following the herd [17], fairness [18], scarcity [19], loss aversion [20], endowment effects [21], how information is framed [22], and the list goes on.

In 2008, the original authors of the book Nudge [1] took many of these insights further by arguing that institutions in the public and private sectors should tailor decision environments to accommodate the constraints of human decision–making [23]. In their book, Thaler and Sunstein [1, p. 9], defined a nudge as "... any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives". Their main premise of nudging was quite simple: In many decision environments, people can make predictable errors of judgment due to sometimes irrational factors. If those errors are anticipated, designers of choice architecture can reduce the error rate [24]. The designer of a particular decision environment, often referred to as the "choice architect" has a significant and often underappreciated influence over what choice people will be making [25]. It should be noted that the general goal of nudging was to make people better off, "as judged by themselves." [1, p. 5]. An innocuous and popular example of a nudge is that of an etched image of a fly inside of men's urinals [24]. This ostensibly small intervention reportedly improves the aim of the target user which, in turn, leads to considerably lower cleaning costs [26].

1.1.2 Nudging in the digital sphere

While Thaler and Sunstein [1] were the first to organize the theoretical principles around nudging and choice architecture - the research concerning the engineering of technology for persuasive effects had already been established prior the publication of their book. Previous scholars have argued that Thaler and Sunstein [1] did not devote too much attention to pervasive computing technology, either as a means for supporting everyday choice or as a context in which choices were being made [6].

Yet, the term 'digital nudging' was eventually adopted within the information systems and computer science communities (see for example [27, 28, 29]), and the term has since gained exponential attention as its own domain of research [30]. The next chapter of the dissertation will show that there are several competing definitions in the literature for what phenomena that could or should fall inside the domain of digital nudging.

In line with the more inclusive definitions on the topic [27, 31], I will view a *digital nudge* as the event where digital artifacts steer users in particular directions while also allowing them to go their own way [30, p. 2]. *Digital choice architecture* will involve the design and configuration of a choice environment - controlling what outcomes are associated with the choice itself, and how these potential outcomes are presented to users via pervasive computing devices.

1.1.3 A larger space of behavioral change interventions

There are numerous domains of research that can be brought to bear on how artifacts can be designed to steer user behaviors. In general, nudging could be fitted within the larger research stream of behavioral change interventions[32]. While a fair and complete outline of behavioral change interventions is too large for the scope of this dissertation – I will briefly echo some prevalent conceptual definitions of the research domains that share significant overlaps with digital nudging. Out of these associated research domains, persuasive technology (PT) arguably has the widest conceptual footprint. Fogg [33] pushed for research in the intersection between computer science and the psychological theories related to persuasion before a coherent idea of nudging was established. PT pertains to "... any interactive computing system designed to change people's attitudes or behaviors" [34, p. 1]. Around 2008, Information Systems scholars established another research agenda for similar phenomena referred to as 'persuasive systems design' or (PSD)[35]. The original proponents of both PT and PSD issued the caveat that these technology-mediated interventions should be performed without coercion or deception [34, 36].

Related to nudging, Thaler and Sunstein [1] highlighted that there are cases where companies have strong incentives to cater to people's potential frailties in order to exploit them. They also remained agnostic as to whether free markets and competition would solve this particular problem. The last overlapping research domain is focused on this problem in the digital domain – which is often called 'dark patterns.' Dark patterns are use cases when the aims or functionalities of a UI are tainted to some or a large extent. Dark patterns began as a more practitioner-oriented idea [37], where Brignull [38] first defined and assembled a collection of such instances. Another early contribution to this domain can be found in Conti and Sobiesk [39]. Mathur et al. [40, p. 1] later defined dark patterns as "... design choices that benefit an online service by coercing, steering, or deceiving users into making decisions that, if fully informed and capable of selecting alternatives, they might not make." Other prevalent definitions also contain some variety of malicious intent on the behalf of the UI designer or other stakeholders [41, 42].

Based on the above definitions, the design spaces of these associated research domains concern similar descriptive phenomena. On an abstract level, they are mainly distinguishable in terms of ethics and underlying motives. Most artifacts can be used for constructive and destructive purposes, and it is clear that interactive technologies can be used for a variety of persuasive aims and this is a very vexing problem. With that being said, it is difficult to descriptively delineate the listed research domains as neat mutually exclusive categories of scientific inquiry. Rather, they are dynamic and continuously deepen and augment each other's research output and will therefore be invoked at relevant points throughout the dissertation.

The main problems that this dissertation addresses related to digital nudges and choice architecture will now be further defined.

1.2 Problem Definition and Justification

The presence of a choice architecture in the digital realm is inevitable, and cannot be wished away. This contention stems from the "inevitability argument" for choice architecture in general from Sunstein [31, p. 4]. Small changes to these choice architectures (or nudges) will have downstream effects on sometimes important outcomes in our evermore digitalized world, and there is currently no obvious path on offer where HCI and systems designers can collectively decide to *completely avoid* influencing the behaviors of their users, by design. Put another way, there are arguably very few online design spaces that allow for a completely neutral choice architecture where the users volitional agency and freedom is entirely unaffected. To illustrate the importance of digital choice architecture, let's consider a few examples related to the application domains that this dissertation will touch. Consider a large car manufacturer that decides to update their Terms of Service (ToS) and Privacy Policy (PP) for their divers and passengers. The Mozilla Foundation recently published a privacy review of 25 major car brands outlining that 76% of these car brands sell their costumers personal data [43]. More concretely, these personal data may include: location -/ health diagnosis data, biometrics, genetic information and sexual activity [43]. Given people's everyday dependency on cars and that many drivers probably consider their car as their own private space, designing meaningful privacy controls for modern cars seems like an ethical emergency. Another example could be a prevalent social media service that decides to populate its users' feeds with more engaging and relevant content by updating its recommendation systems. Several studies have indicated that content focused on moral outrage or out–group animosity is shared more often and spread more rapidly on social media platforms [44, 45].

As a final example, consider when a popular online donation platform decides to change the default design of how projects are presented and ranked. Previous research has demonstrated that various design elements matter in the context of large-scale donation behaviors [46]. By using snapshots from 'The Internet Archive' of DonorsChoose.org together with longitudinal donation data from the site itself, Chakraborty et al. [46] demonstrated how ostensibly minor features such as default rankings impacted the distribution of users' donations*. Some of the businesses in the above examples do operate under quite peculiar incentives. Davenport and Beck [49] refer to these incentives as the attention economy, meaning that users' attention is a very valuable currency and managing attention is perhaps the key determinant of business success in our digitized societies. This becomes especially relevant when online ads serve as the main source of revenue for a business.

While business incentives definitely matter, the finer point of the above examples is that ostensibly small design elements, such as search input fields, various lists or feeds, menus or 'I Accept' buttons constitute a choice architecture in various situations for users. These design elements may, in turn, exert various forms of influence on users. The digital sphere also provides different stakeholders (or choice architects) with more capabilities to readily change various designs coupled with the tools for measuring the impacts of such changes. This further raises the stakes and potential consequences for nudging online users in various directions. Given its importance, previous scholars have argued that an evermore pressing but under-researched area concerns digital choice architecture[50].

*Similar findings have been found for design features related to search engines (such as default rankings and auto-completed search fields) and their potential impact on democratic elections [47, 48].

In short, the chapters of the dissertation will explore how digital nudges can be effectively and ethically utilized to positively influence user behaviors and attitudes in a set of online environments. We will also investigate their potential adverse effects on online users. This area can be further broken down into four over-arching questions, each of which will be further specified and addressed in the subsequent chapters of the dissertation:

Q1: What is the current state of digital nudging research, and what are the existing gaps related to the empirical evaluation of digital nudges?

Q2: How can digital nudges be effectively utilized to increase privacy awareness of users interacting with popular online services?

Q3: How will different default newsfeed settings influence user behavior and experience on a popular social media platforms?

Q4: How do combinations of digital nudges impact users' charitable giving behaviors, and what are the resulting effects on users' ethical concerns towards these designs?

1.3 Contributions

The dissertation advances the understanding of digital nudges and choice architecture by investigating their diverse impacts (or sometimes lack thereof) across several application domains. We provide various starting points for future research by highlighting unexplored areas and potential approaches to address the design challenges associated with digital nudging. The user-studies in this dissertation reveal insights about how design and system scholars may facilitate users' privacy awareness, moderate compulsive use of social media, or combine charitable giving features online.

1.3.1 Outlining open research avenues for digital nudging

The first contribution provides a systematic overview of existing literature on digital nudging, showcasing a ten-fold increase in research output for the domain over a five-year period. We also identify current research gaps around how these design interventions are empirically evaluated and highlight open research questions. The review spans seven scientific databases from which we analyze 109 separate empirical studies on digital nudging. The analysis from our literature review lays out nine possible paths for future studies in the field of digital nudging and choice architecture.

1.3.2 Augmenting online privacy with digital nudges

The second contribution provides a relevant demonstration of how PP and ToS agreements could be made more accessible to Internet users with digital nudges. Through a double-blind online experiment ($n = 183$) and a descriptive field study with users from the IKEA Place app ($n = 81431$) we show how to increase users informed consent to the ToS and PP of popular online services by employing subtle changes to their current digital choice architecture.

1.3.3 Constraining a newsfeed for social media detox

The third contribution concerns the mitigation of compulsive social media usage, specifically on Facebook. Previous research suggests that the infinite newsfeed is often associated with a sense of regret according to users themselves [51]. Through an online field experiment ($n = 138$) we change the default design mechanics and thereby create different newsfeed diets for a sample of real Facebook users. This study provides a demonstration for moderating users' time spent on social media. We also provide an in-depth evaluation of these design interventions related to user experiences.

1.3.4 Combining pro-social nudges for online charitable giving

The fourth contribution contains a detailed account of how combinations of three common digital nudges statistically interact when they are deployed in a charitable giving context. In an online experiment ($n = 794$) with real monetary stakes for users, we design and deploy three digital nudges, and measure the impact of their various combinations on users' donation behaviors and attitudes towards these designs. We also investigate two important ethical dimensions in relation to nudging, namely autonomy and manipulation [50], but from users' own perspectives. Our evaluation reveals potentially significant costs in terms of users' sense of autonomy and perceived manipulation.

1.4 Outline and Composition

The dissertation is structured as a set of chapters. This means that each chapter may be read as an independent article but it also creates repetitions of some key concepts throughout the dissertation. The research in the dissertation has been presented in a number of peer-reviewed articles and a working draft paper. Versions of chapter two to four have appeared in international publication venues in the broader research areas of information systems and computer science.

1.4.1 Publications

The research featured in chapter two to five and has been disseminated in the following past or upcoming publications, see Table 1.1. The first page of every chapter indicates whether the contents have appeared in one of the listed publications and provides a more detailed outline of the chapter's contents.

Table 1.1: Mapping of the chapters and prior or future publications in the dissertation.

Chapter	Based on the following prior or upcoming publication
2	Kristoffer Bergram, Marija Djokovic, Valéry Bezençon, & Adrian Holzer. *The Digital Landscape of Nudging: A Systematic Literature Review of Empirical Research on Digital Nudges.* 2022. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22). ACM, New York, NY, USA, Article 62, pp. 1–16. [30]
3	Kristoffer Bergram, Valéry Bezençon, Paul Maingot, Tony Gjerlufsen & Adrian Holzer. *Digital Nudges for Privacy Awareness: From consent to informed consent?.* 2020. In Proceedings of the European Conference on Information Systems (ECIS '20). AIS, Marrakech, Morocco, pp. 1–16. [52]
4	Aditya Purohit ‡, Kristoffer Bergram ‡, Louis Barclay, Valéry Bezençon & Adrian Holzer. *Starving the Newsfeed for Social Media Detox: Effects of Strict and Self-regulated Facebook Newsfeed Diets.* 2023. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). ACM, New York, NY, USA, Article 196, pp. 1–16. [53]
5	Kristoffer Bergram, Abdessalam Ouaazki, Manon Berney, Valéry Bezençon, & Adrian Holzer. *Stacking Nudges for Online Charitable Giving: An exploratory analysis of pro-social nudge combinations and their polarizing effects from a user's perspective.* A working draft, pp. 1–21. [54]

‡ *Both authors contributed equally to this work.*

1.4.2 Structure

The rest of this dissertation is structured in the following manner: Chapter two reviews the literature on digital nudging across several relevant dimensions such as how they tend to be designed (their design patterns), how their impact tends to be evaluated in terms of methods and user samples, and maps whether they belong to a personalized and/or interconnected choice architecture.

The third chapter explores how small changes to the digital choice architecture for ToS and PP consent notices can be used to increase the privacy awareness of users. Through an online experiment and a descriptive field study we show how phrasing this agreement differently and providing a highlights alternative to the standard UI designs can move users closer towards informed consent.

Chapter four is an detailed analysis of compulsive social media use. We investigate the differences in terms of time spent and users' self-reported experiences between a control group (where no changes are made to a social media newsfeed), a strict newsfeed diet (where the newsfeed is automatically reduced to a minimum), and self-regulated newsfeed diet (where the newsfeed is reduced, but users can then re-follow their contacts).

Chapter five explores systematic combinations of default, friction and social nudges to understand their additive and interactive effects on users' charitable giving decisions. We provide an in-depth evaluation of users' own perspectives on how these nudge combinations relate to their perceived threats to autonomy, sense of manipulation and how users themselves describe these nudge combinations.

Chapter six provides concluding remarks by summarizing the research contributions and discussing future research in the domain of digital nudging and choice architecture.

Chapter seven lists all the references for the dissertation. Since the dissertation is structured in a connected article format, the references are indexed on each individual chapter. While references may appear across several chapters, references will listed and numbered under the chapter where they first appeared in the dissertation. The final chapter contains an appendices with auxiliary information for some of the analyzes in the other chapters of the dissertation.

2 Digital Landscape of Nudging and Choice Architecture

A version of this chapter has been published as: **Kristoffer Bergram**, Marija Djokovic, Valéry Bezençon, & Adrian Holzer. *The Digital Landscape of Nudging: A Systematic Literature Review of Empirical Research on Digital Nudges.* 2022. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22). ACM, New York, NY, USA.

Contents

2.1 Introduction

As we saw in the previous chapter, the concept of *nudging* emerged in the field of behavioral economics and policy-making around 2008 with the publication of the book "Nudge" [1]. It has gained momentum since then to become one of the most transformational trends in the public sector during 2020 [55]. Undergirding the concept of nudging is the concept of choice architecture: an environment where choices and behaviors occur. These choice architectures are omnipresent and range from natural phenomena such as the weather, to abstract concepts such as policies, or concrete objects such as furniture designs [56]. A concrete example to sensitize people to both the opportunities and the challenges that come with the concept of choice architecture is the arrangement and display of foods in a cafeteria [57, 1]. This example shows that there is no neutral way to arrange and present the foods in a cafeteria leading the designer to inevitably guide the choices and behaviors of guests.

In the literature, there is a lively debate on where to set boundaries as to what qualifies as a nudge. Definitions range from general to specific (see [31, 58, 59, 60] for a review). Among the general definitions available, a nudge could be broadly defined as an intervention that steers people in particular directions but that also allows them to go their own way [31, p. 417], or as "an intervention on the choice architecture that is predictably behavior-steering, but preserves the choice-set and is (at least) substantially non-controlling, and does not significantly change the economic incentives" [61, p. 343]. Nudges can be differentiated from changes to incentive structures or pure mandates [1]. More specific definitions start with a similar description but add some further requirements. For instance, Hansen's definition [58, p.164] includes requirements on the underlying psychological mechanisms used and proposes that "A nudge is any aspect of the choice architecture that alters people's behavior in a predictable way without (1) forbidding any options or (2) significantly changing incentives, whether regarded in terms of time, trouble, social sanctions, economic and so forth. (3) Nudges are called for because of cognitive boundaries, biases, routines, and habits in individual and social decision-making, and (4) work by making use of those boundaries, biases, routines, and habits as integral parts of the choice architecture." Further definitions also include ethical criteria in their requirements, such as transparency, welfare, respect, or autonomy of the nudgee (e.g., [62, 63, 60, 61]). If these criteria are not met, the change in the choice architecture can be called a sludge or a nag instead of a nudge [64].

2.1.1 Digital nudging

With the proliferation of technology, the notion of digital nudging has started to appear with several definitions derived from the literature on traditional physical nudges. General definitions include adaptations of Sunstein's definitions of nudges in the physical realm, typically specifying that the intervention is digital or happens in a digital environment such as "the use of user-interface design elements to guide people's behavior in digital choice environments" [27, p. 433]. Specific definitions start with a similar definition and then list certain criteria (descriptive and/or normative) that a digital intervention needs to meet in order to be considered a digital nudge. For instance, Lembcke et al. [60, p.10] start their definition of the digital nudge as "...any intended and goal-oriented intervention element (e.g. design, information or interaction elements) in digital or blended environments attempting to influence people's judgment, choice, or behavior in a predictable way", then they further require such interventions to be made possible by, and to make use of, certain cognitive mechanisms (e.g., biases, routine), preserving full freedom of choice, being transparent, and increasing "the private welfare of the nudged individual (pro-self) or the social welfare in general (pro-social)". In this paper, we opt for an inclusive definition of digital nudge in line with [27, 31], as the event where digital artifacts steer people in particular directions while also allowing them to go their own way.

There are two characteristics to consider in relation to choice architecture that differentiate digital nudging from conventional nudging: *personalization* and *interconnectedness*. First, a combination of more personal and diverse delivery systems (mobile [65], ambient [66], wearable [67]) and the increased capacity for data analytics provides unprecedented capacity to (1) gather and use real-time data (e.g., location, user demographics, user actions) in order to (2) dynamically personalize the choice architecture of individual users. Second, the ubiquity of social media services not only presents users with personalized choice architectures, but also interconnects the choice architectures of one user with those of others. That is, (1) a choice architecture of a user can contain information from other users and (2) the actions and choices of a user can, in turn, dynamically modify the choice architecture of other users. This type of environment makes it very difficult to predict the outcome of a change to the choice architecture (i.e., a nudge), since it does not operate in a vacuum. These characteristics make it particularly important to understand the consequences of digital nudging design, as these might be unintended and far-reaching.

Returning to the cafeteria example: the cafeteria or any physical environment has certain native constraints. It is very challenging to personalize the arrangement and display of foods for different guests. Also, the choice architecture of the cafeteria is not really interconnected between guests, except in a very local sense. The choices and actions of guests will not immediately spill over into another cafeteria and modify its arrangement and display of foods. Conventional nudging is therefore mainly restricted to non-personalized and disconnected choice architectures. This contrasts with digital services which can unlock tremendous value for users with personalized and interconnected choice architectures. However, it is also these two aspects that may create very unexpected externalities that often accrue to users or society as a whole. This includes far-reaching consequences such as the algorithmic influence of democratic elections [68], the rapid spreading of fake news in online environments [69], the inadvertent radicalization of online users [70, 71], or the tendency that a great majority of users agree to terms and data privacy policies without ever reading them [72].

2.1.2 Current knowledge

While the above problems may be far outside the normal scope of HCI designers, implementing potential solutions will inevitably fall into the lap of the HCI community. These problems are all connected to the digital choice architecture of several online services. While none of these problems have easy solutions, more systematized knowledge on the topic of digital nudging can offer more conceptual clarity to address these problems from a design perspective. Researchers have started tackling a diverse set of issues, as digital nudging has become a hot topic in academia with a ten-fold increase of the literature over the last five years (see Figure 2.1). Furthermore, several well-researched review articles have emerged in recent years to shed light on various aspects of nudging and digital nudging, such as effectiveness [73], categorization [74], psychological underpinning [75, 76], or analyses of their application in specific areas (e.g., privacy/security [76], recommender systems [77]). For instance, Hummel et al. [73] conducted a quantitative review of the relative effect sizes of nudges. In their paper, they categorize nudges into ten different categories (default, simplification, social reference, change effort, disclosure, warnings/graphics, precommitment, reminders, elicit implementation intentions, feedback). Their findings suggest that, based on median effects sizes, defaults seem to be the most effective type of nudge, while commitment seems to be the least effective.

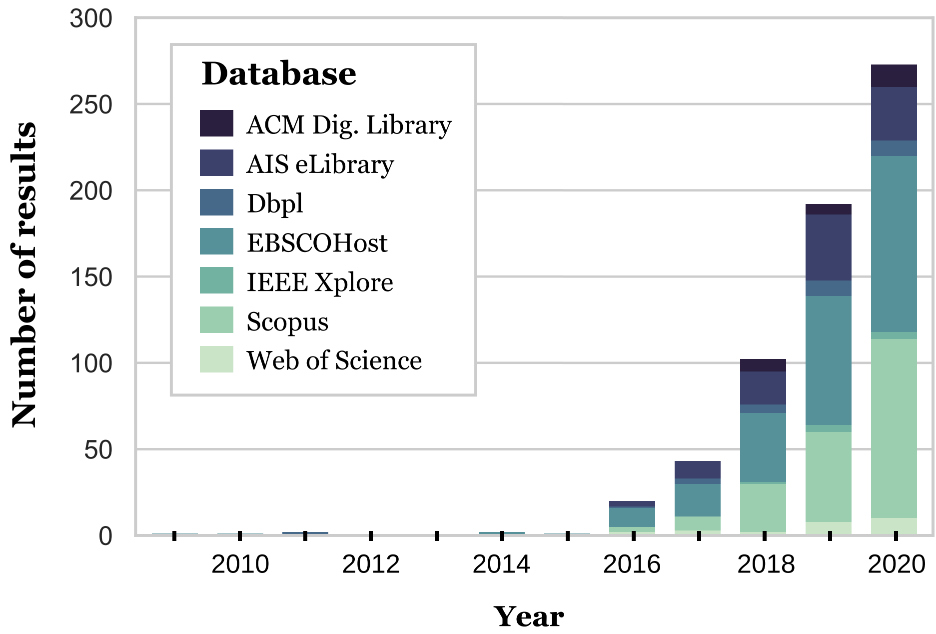

Figure 2.1: The number of search results of the term "digital nudg*" in seven databases between the years 2010–2020

It should be noted that even though Hummel et al. focused on conventional nudges, they also investigated examples of digital nudges and observed that their effects did not differ significantly compared to conventional nudges. This finding can be put in perspective with the research conducted by Caraban et al.[74]. In their study, the authors provided a review of digital nudges and synthesized them in 23 different mechanisms, grouped in six high-level patterns: facilitate, confront, deceive, social influence, fear, and reinforce. Even though they do not specifically focus on empirical studies, they still provide a rich discussion of the evaluation results found in the literature. For instance, they observe that most (66%) nudges evaluated were effective. They also discuss several pitfalls of nudges (e.g., lack of educational effects, no long-term effects, backfiring) and argue that even though personalization has been highlighted as an important aspect of digital nudging [74, 77], it has not yet been investigated systematically and thus presents an "untapped opportunity" [74, p. 11].

2.1.3 Purpose and aims

While the interest in digital nudging has increased and several reviews have provided valuable insights into this young field, the landscape of digital nudges has not been fully mapped. Previous literature has provided valuable insights into digital nudge mechanisms. They pointed to potential positive effects of nudging, and pitfalls, and discussed the underlying psychological mechanisms. However, there are several areas of digital nudging that warrant further systematic investigation:

(1) Previous reviews have not mapped how the literature has investigated the specificity of digital nudges that we highlighted above: personalization and interconnectedness.

(2) While digital nudges are increasingly employed by researchers and practitioners – the specific problems that they address have not been sufficiently mapped across the literature.

(3) Digital nudges can be combined in a variety of ways, and evaluated using different methods and user samples. These practices have not yet been mapped out in the literature.

In this chapter, we describe and synthesize empirical research contributions in the digital nudging literature across HCI and other related scientific fields. We classify these contributions across 10 established digital nudging designs, referred to as "patterns". We then explore this design landscape by mapping these patterns across the three areas that we have highlighted above. First, we explore this design landscape by mapping these patterns across two characteristics of choice architecture, and their current delivery channels/devices. Second, we explore the problems addressed by these nudging patterns by outlining their application domains and targeted outcomes in terms of behaviors and attitudes. Third, we investigate the evaluations of digital nudging solutions by mapping their evaluation methods, sample types, and highlight whether combinations of nudging patterns have been evaluated in the literature. Finally, building on the unexplored patches in the digital nudging landscape, we contribute to the digital nudging literature by setting a future research agenda organized around nine high-level research questions to be investigated by the HCI-community.

2.2 Methodology

The methodology of this literature review will be described in four phases based on recommendations for systematic reviews [78]: literature search, screening filter, eligibility evaluation, and analysis procedure.

2.2.1 Literature search

The goal of the literature review was to survey the landscape of research that explicitly tackles digital nudging, as opposed to the broader research that can be considered as a digital nudge but is not presented as such by the authors. To capture as many contributions as possible, we included seven major databases spanning a wide range of research fields from HCI, computer science, and information systems to economics and psychology. The following databases were searched: ACM Digital Library, AIS eLibrary, Dbpl, IEEE Xplore, EBSCOHost, Scopus, and Web of Science. We identified potential papers that authors characterize as digital nudging, using the following strings across the selected databases: *"digital nudging"* OR *"digital nudge"* OR *"digital nudges"*. All databases were accessed on 9th January 2021, yielding 638 results in total. The PRISMA flowchart in Figure 2.2 is adopted and modified from Liberati et al. [79]. Following the recommendations by vom Brocke et al., Figure 2.2 also highlights the functionalities used, and the number of identified papers from each database [80]. We filtered the publication period for all the results in this review between 2008–2020.

2.2.2 Screening filter

Two of the co-authors deleted duplicate records which resulted in a dataset containing 361 unique results. Very incomplete entries were removed, e.g., papers lacking a title, publication year, and publication venue. Entries with titles that were not written in

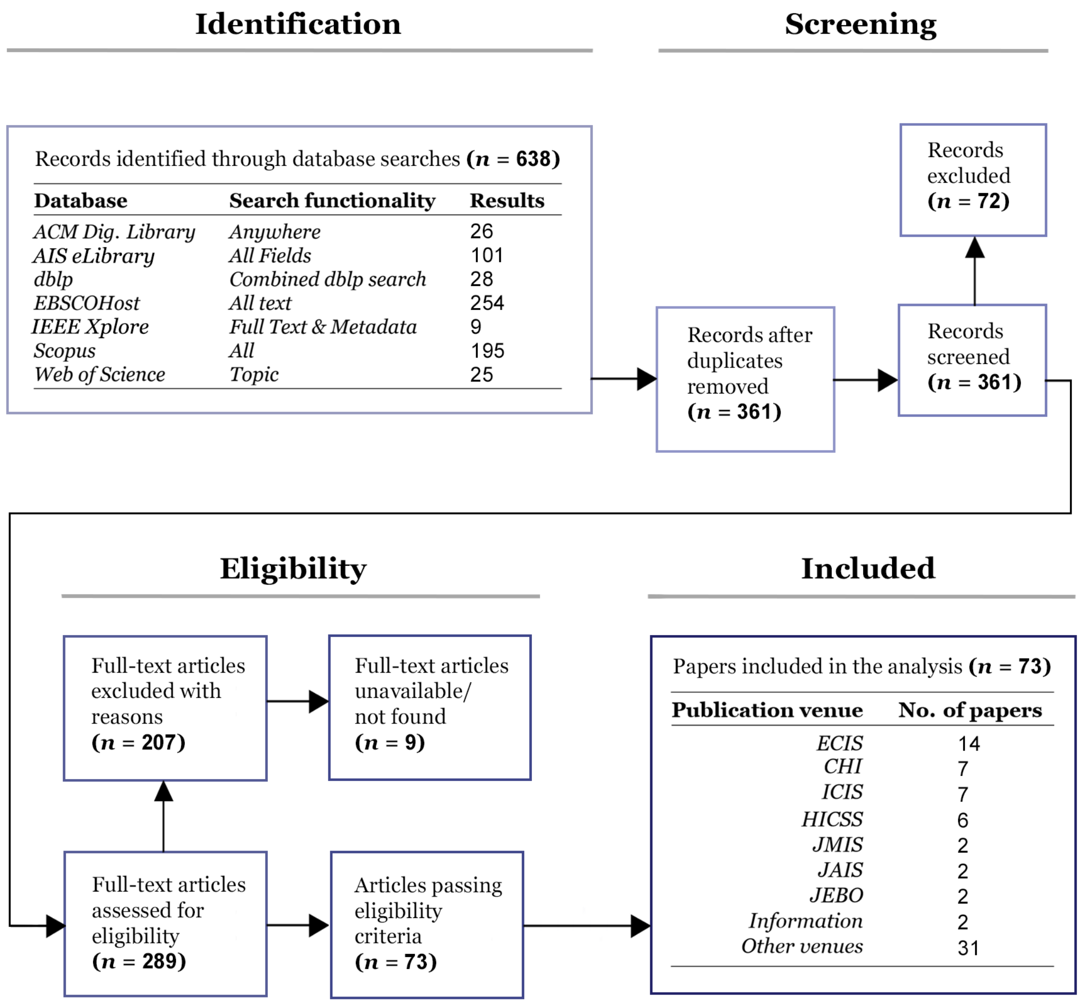

Figure 2.2: Overview of literature review methodology (adapted and modified PRISMA flowchart from Liberati et al. [79])

the English language were removed. As in other reviews on similar topics - entries that were not published in a peer-reviewed conference proceeding or journal were removed as a proxy for quality control [74, 76, 81]. This last step removed entries such as proceeding adjuncts, workshops, dissertations, student consortiums, or talk papers, leaving 289 full-text articles.

2.2.3 Eligibility evaluation

During the third stage of the review, the papers were distributed among three of the co-authors to ensure that each paper had a full-text assessment by at least two reviewers. The following two criteria were specified to include appropriate articles for the review: the paper had to (1) mention *"digital nudge"*, *"digital nudges"*, or *"digital nudging"* in the full-text, excluding the references, and (2) the outcomes of the digital nudge had to be observed or tested to some context, i.e., empirically demonstrated/evaluated. Conflicting assessments on the criteria were noted and resolved by at least two reviewers.*. Three papers were excluded because they relied on the same empirical studies as two other included papers. In these cases, we included the most recent iteration of the papers in the final analysis. Following the example of Connolly et al., a random sub-sample of 59 articles (20 %) was assessed by an independent reviewer and compared to one of the co-authors to gauge inter-rater reliability [89]. All eligibility criteria between these two assessments showed good agreement overall (*Cohen's Kappa* = 0.74). Finally, 73 papers met the eligibility criteria and were included in the analysis. Figure 2.2 provides an overview of the process.

These papers were included though they only referenced "digital nudg" without explicitly mentioning it in the full text: [82, 83, 84, 85, 65, 86, 87, 88]

2.2.4 Analysis procedure

To synthesize and analyze the included literature, a concept-centric approach was used in accordance with Webster and Watson [90]. A concept matrix was compiled in an iterative fashion while the included papers were read by the co-authors. Primarily, two of the co-authors coded and classified the selected articles with the assistance of a third. When codes or classifications were difficult to infer for a specific paper, we discussed these issues until we reached agreement.

First, highly accessible and descriptive information was extracted, such as publication year, authors, publication venue, number of participants, and application context. Next, we coded detailed information about the digital nudges being investigated in the papers along the following dimensions: type of outcome, nudge pattern, choice architecture, delivery devices, delivery channels, study's sample type, and evaluation methods. Depending on the coded dimension, the unit of analysis varied between papers, studies, and digital nudges. Indeed, one paper may contain several studies investigating different nudge patterns, outcomes, or different evaluation methods.

2.3 Digital Nudging Landscape

The analysis of the digital nudging landscape is based on a diverse sample of 73 peer-reviewed papers. While Figure 2.2 shows that the the majority of included papers are clustered in well-known publication outlets within information systems and computer science, the whole sample of papers extends across 38 unique conference and journal venues spanning several scientific fields. The included papers contained 109 discrete studies in total. The number of observations in all these studies ranged from 10 to over 1 million ($median$ = 183). Across these studies, we identified 231 separate digital nudging interventions that were demonstrated and evaluated.

Table 2.1 provides a comprehensive overview of all papers and their codes included in the literature review.

2.3.1 Design of digital nudges

We analyzed the design aspect of digital nudges in the literature from the perspective of the nudge patterns, the characteristics of the surrounding choice architecture, and the delivery channels/devices.

Digital nudge patterns

To steer users in a certain direction, different nudge patterns can be chosen. These patterns, or design solutions to a recurring problem [91], are sometimes referred to simply as nudges [92], nudge categories [73], or nudging mechanisms [74]. We drew on previous literature [92, 73, 74] to assemble a list of the ten main nudge patterns: social, reinforce, disclosure, friction, feedback, default, warning, scarcity, deception, and commitment. We settled on these patterns after recursively labeling and comparing our corpus of papers with patterns in the literature. More specifically, we derived seven patterns from the list of ten nudge patterns presented by Hummel et al. [73], and Sunstein [92], grouping together certain categories into a single pattern (e.g., change effort and simplification into friction) because they were difficult to differentiate during coding. To these we added three patterns derived from Caraban et al. [74], who categorized technology-mediated nudges specifically, and which were difficult to classify as other patterns: reinforce, scarcity, and deception. Figure 2.3 gives an overview of the digital nudge patterns among the analyzed papers. Below, we provide a brief summary of each pattern with concrete examples of its application from the literature.

Social nudges. This nudge guides the user's behavior by providing references to how other users behave, thereby creating a social norm [92]. In a classical example, researchers investigate the influence of a social nudge for towel reuse in a hotel [93]. In the hotel room, a sign tells a group of guests they can reuse their towels if they wish to help the environment. Another group in the social nudge condition sees a similar sign with a text mentioning that almost 75% of guests participate in the program[92]. The researchers show that towel reuse is significantly increased in the latter group [93]. In the reviewed literature, these social nudges are often used to make it appear as if the choice architecture is interconnected without it actually being so. To illustrate: Kretzer and Maedche [86] investigated how social nudges can be used in the context of business information systems (BIS) to encourage employees to reuse reports. More specifically, they evaluated if employees would tend to reuse a report more if there was a mention next to it indicating that someone else with similar characteristics liked the report, e.g., "Ian, a project leader from the accounting department in France, liked this report." They found that users were more inclined to reuse a report if it was liked by someone from a similar rank (intern vs director), belonging to the same department or coming from the same location. In an online experiment, DiCosola and Neff showed how different forms of social comparisons could nudge users to reduce the number of calories contained in their shopping baskets in an e-commerce scenario [94]. By conducting their study on a sample of healthy-weight adults, out-group comparisons were achieved by providing calorie benchmarks for overweight adults. Surprisingly, out-group comparisons seemed to work just as well if not better than in-group comparisons for eliciting more healthy choices.

Reinforcement nudges. These nudges reinforce behaviors and choices by increasing their salience in the mind of the user [74]. This entails practices that rely solely on persistent or frequent exposure, or some underlying psychological mechanism such as priming, where a stimulus is associated with a desired behavior (when asked to complete the word "so_p", participants previously primed with a picture of a shower tend to say "soap", whereas participants primed with a picture of food, tend to say "soup" [95]), or anchoring where potentially unrelated information influences the outcome (telling someone Mark Twain's birth year can influence their estimate of the length of the Mississippi river [95, 96]). These nudging patterns are perhaps the broadest in this literature review. As an illustration, Dennis et al. [97] conducted seven lab experiments to investigate how both numeric and semantic priming could be used to sway users' willingness to pay in e-commerce/marketing contexts. In their paper, priming meant that users were exposed to a stimulus (a product and its price) that was meant to later affect their intention to make certain purchases. Their experiments highlight that both kinds of priming seem to work better in online auction settings than in normal e-commerce shops. In the former, it is possible to sway users to pay more for products through priming since the value of auction products is often uncertain. Reinforcement nudges can also be applied to problems in education. On a large sample of university students, Brown et al. [98] showed how a combination of reinforcement and social nudges in a web-enabled coaching system could increase students' proactivity to get started on their homework.

Disclosure nudges. As a nudge, disclosure entails adding information that is accessible, clear, and relevant to the choice that the user is about to make [92]. Product labels indicating energy efficiency are typical examples of disclosure nudges (e.g., [99]). In the reviewed articles, Gimpel et al. [100] investigate how a disclosure nudge can increase the ability of social media users to identify fake news. In their experiment they show participants a series of articles posted by news providers (some truthful, some not). In the nudge condition they disclose related articles below the main article to provide additional knowledge to users. They find that this nudge works best if the related articles are a mix of truthful and fake news articles. Indeed, if the related articles are only truthful, the authors find no increase in fake news detection.

Also, various kinds of disclosure nudges can sometimes have harmful consequences. In the UI of ridesharing platforms, Abramova [87] demonstrated through an online experiment how different disclosures (such as a Middle Eastern male name in the profile of the driver) decreased West European users' willingness to pay for a ride.

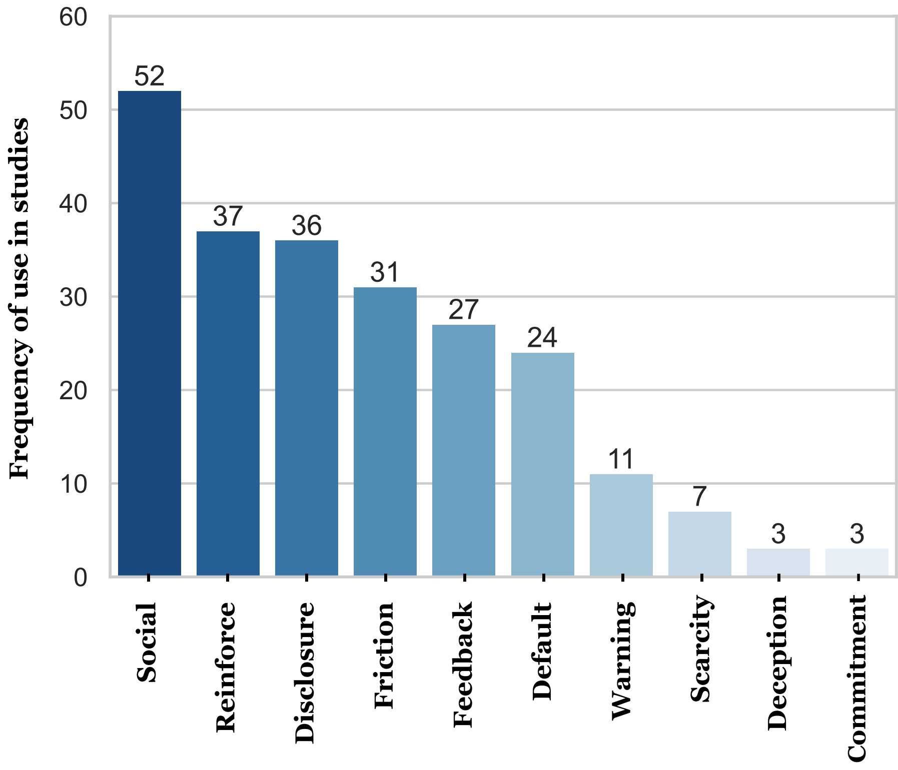

Figure 2.3: Number of empirically evaluated digital nudges in the papers, across the ten patterns (unit of analysis: nudges)

Friction nudges. A behavior can be encouraged or discouraged by removing or adding friction respectively. Sunstein [92, p. 4] refers to this nudge as "increases in ease and convenience", Hummel et al. [73, p. 52] uses the label "change effort", and Caraban et al. [74, p. 6] identified a mechanism they called "creating friction". We have adopted the one-word descriptor above to convey that this digital nudge pattern can be used in two ways: by removing or by adding friction to the user experience. Furthermore, while simplification is sometimes thought of as its own type of nudge – we see it as a means of removing friction for the user. Kim et al. [65] showed in a field experiment how a lockout task with zero to very small workloads could limit smartphone usage. Their paper demonstrated how the simple friction of an app arbitrarily asking users to input 0–30 random numbers helped users to self-regulate their smartphone app use over time. Friction nudges can also provide valuable solutions in the context of policy-making. In an extensive analysis of field data, Fox et al. [101] evaluated digital nudges targeting the friction of accessing healthcare in the US. Online features such as receiving real-time eligibility decisions, or opting in for presumptive eligibility had considerable effects on enrollments over time. This was especially true for children since these nudges remove the administrative friction that accrues to parents.

Feedback nudges. Whereas a disclosure nudge provides information about an upcoming choice of a user, a feedback nudge provides information about a past or a current behavior of a user. An example of such a pattern in the reviewed papers include a field experiment where Hoffmann and Thommes [66] applied feedback nudges to a fleet of truck drivers. They increased truck-drivers' eco-efficient driving behavior by using an eco-score shown on the truck's telematics system. Okeke et al. [102] designed and evaluated a mobile app that delivered haptic feedback in the form of continuous vibration if users exceeded their daily Facebook limits.

Default nudges. Defaults are perhaps the most well-known and most effective pattern for nudging [92, 73]. The idea of a default nudge is to design the choice architecture in such a way that the default behavior is the desired behavior [103]. The classical example of this nudge pattern is organ donation enrollment policy. The policy can dramatically increase the percentage of people consenting to being donors by changing from a default set to non-donor with a possibility to explicitly opt in, to a default set to donor with a possibility to opt out [104]. While also investigating two specific personality dimensions and their relationship to nudging in an online experiment, Ingendahl et al. [105] changed the pre-selection of various product categories in an e-commerce scenario. While the measured personality dimensions showed little to no detectable relationship to nudging, the default nudge was successful in altering product choices among users. This was especially true when the default was combined with a social nudge[105]. In a later section we will look at how frequently different nudging patterns were combined and investigated together in the literature.

Warning nudges. Different kinds of warnings and graphics assist users by grabbing their attention to various risks or consequences [92]. This is important because digital choice environments tend to abstract real-world outcomes for the user. In a lab experiment designed to investigate decision inertia with financial planning tools, Jung and Weinhardt [106] demonstrated how warning nudges from a robo-adviser can mitigate this problem even when controlling for financial literacy among users. Yet, their study also suggested that default nudges were more effective in this context. Another study demonstrated how warning nudges could help protect user privacy by raising potential unwanted data disclosure on social media [107]. To illustrate, when a user is about to post about himself and his friend Yvette having dinner at Brotzeit despite a headache, a warning message would point to sensitive information that the system could gather from this post, e.g., "potential sensitive topic found (health), potential identifier found (Yvette), potential location found (Brotzeit)" [107, p. 4431].

Scarcity nudges. Scarcity nudges work on the assumption that people assign more value to something that is going to be more difficult to acquire in the future [108]. Typical examples of scarcity nudges found online include the indication that there is "only one room left" on hotel booking platforms [109]. Similarly, in the reviewed papers, Schneider et al. [110] used a scarcity nudge to increase the selection of a crowdfunding option by placing the indication "only five left" next to it. Their results show that there is a reversal in preference towards the option with the scarcity nudge. Another example is provided by Wessel et al. [111] who examined how a scarcity nudge in the form of sold-out early birds influenced users' behavior in crowdfunding campaigns. Early bird reward options are often at a discount to encourage early backers of crowdfunding campaigns. Yet, even when they were unavailable as a choice but still visible to the user, they nudged users to choose the comparable but undiscounted reward options more often.

Deception nudges. This pattern uses mechanisms that covertly affect how choice alternatives are perceived by users [74]. A typical example of a deception nudge is adding a decoy option to the choice architecture. A decoy option is an option that will not be selected by the user, but that steers their choice to one specific alternative. Schneider et al. [110] provide a good illustration of this nudge pattern. In the context of a crowdfunding platform, they offer two options to users: (A) pay $10 get the eBook, (B) pay $20 get the eBook and the hardcopy of the book. In this setting, it is not easy to compare the option. In the end 69% of participants choose A. Add the decoy option (C): pay $20 get the hardcopy of the book – and the result is the reverse: 68% of participants choose B. Predictably almost no one (1% of participants) chooses C as it is objectively worse than B, but the fact that it allows to set B as objectively better than one option increases its appeal.

Commitment nudges. The idea of this pattern is that when commitments are elicited from the user, they will motivate the user to behave in a way that is consistent with those previous commitments [112]. In a small online experiment, Kroll et al. used commitment and social nudges in a mockup UI of a smart home app to encourage users to save more energy [113]. The researchers elicited a commitment from users by letting them preselect from what household devices (e.g., dishwasher, lightbulbs etc.) they wanted to save energy. They detected no impact on energy consumption attitudes among users when these digital nudges were used on their own. However, a combination of commitment and social nudges led to more energy-saving choices. In a later section we will highlight how frequently such combinations of nudging patterns were investigated in the literature.

Choice architecture characteristics

We coded each nudge investigated in the literature based on the characteristics of their choice architecture in terms of interconnectedness and personalization.

Interconnectedness. The interconnectedness dimension has three levels: Level 0 none, Level 1 partial, Level 2 full. Partial interconnectedness is met when a study investigates how information from other users affects user behavior in a choice architecture. Full interconnectedness is met when the study also investigates how actions of one user, in turn dynamically modify the choice architecture of other users. A rare example of a study that reaches the second level of interconnectedness is a paper by Antinyan et al. that examines the effect of a social nudge, which additionally includes either a disclosure nudge or an incentive (a tax), to decrease the consumption of positional goods (e.g., luxuries) compared to private goods (e.g., necessities) [114]. The study takes the form of a game played over 30 rounds by four interconnected participants. At every round participants can spend a certain amount of virtual money on either private goods or more expensive positional goods. Before making their decision they receive information about how much the others consumed during the previous round. As such, their choice architectures are fully interconnected, as (1) they receive information about the behaviors of others (a social nudge) and (2) their behavior, in turn, will influence the choice architecture of others in the following round. The game rules are set so that everyone would benefit if the consumption of positioning goods was low, but the rules also push individuals to consume more positional goods than the others, which results in excessive positional goods consumption and substantial welfare loss. In the disclosure nudge condition, participants receive information about the negative consequences of excessive positional consumption, whereas in the tax condition, a 25% consumption tax is introduced after a certain level of positional goods consumption. The results show that both interventions work well only when the choice architecture is interconnected and all participants are aware of each other's behaviors. Studies that reach the first level of interconnectedness are more common. Indeed, studies examining social nudges do by definition integrate information from other users in the choice architecture. Schneider et al. provide a representative example of such studies. They used a social nudge to encourage adoption of an electronic identification (eID) by showing participants that 77% of previous users had opted for it too [115]. However, the behavior of the participant does not in turn spill over to change the choice architecture and potentially the behavior of others.

Personalization. The personalization dimension also has three levels: Level 0 none, Level 1 partial, Level 2 full. The first level is met when a study gathers user data (e.g., location, user demographics, user actions) in order to infer the potential influence of the nudge on user behavior. The second level is met when such information is used to dynamically personalize the choice architecture of individual users. A rare study

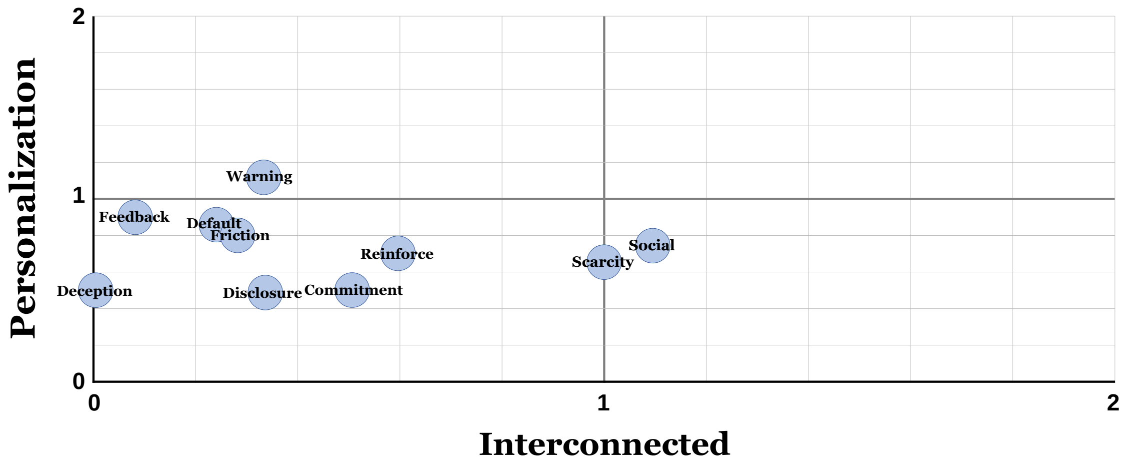

Figure 2.4: Mean level of interconnectedness and mean level of personalization per digital nudge pattern (unit of analysis: studies)

that reaches full personalization used the smartphone's vibration feature as a feedback nudge for users to spend less time in the Facebook app [102]. For one of the study's conditions, the trigger for this vibration was personalized, based on users' past daily visits and their time spent in the Facebook app. If users in this condition reached 50% of their past daily visits or time spent (this threshold was derived from a baseline week) then their phone started vibrating every five seconds. As one might expect, users in this condition spent significantly less time on Facebook. As with interconnectedness, studies that reach the first level of personalization are more common and a typical example can be found in Buck et al. [116]. In a series of experiments using an outcome measure called the app information privacy concern (AIPC), they show how privacy attitudes of users are sensitive to various patterns of nudges, and associated to the user's gender or the operating system (OS) of the user's smartphone. While examples can be found on both interconnected and personalized choice architecture, most research on digital nudging is conducted with the user's choice architecture being disconnected and non-personalized. This result is shown in Figure 2.4 where most of the nudge patterns end up in the bottom left quadrant.

Delivery channels and devices

We also investigated the delivery channels (visual, audio, haptic) and delivery devices (e.g., desktop, mobile, wearable, ambient) of nudges. Our results show that the delivery channel for these digital nudges is almost exclusively visual, that is, via a screen. The reviewed literature only contains two exceptions, which are haptic [117, 102]. The first haptic study examines how feedback can be provided through nerve stimulations [117]. In that study, the goal was to help pedestrians with navigation. As participants would approach a crossroad where they needed to turn left, the system would deliver a nerve stimulation through a wearable electrode placed on the left arm causing the arm to move. The other example comes from the previously mentioned study by Okeke et al. [102], where the continuous vibration of a smartphone would be triggered if users exceeded their daily time limits on Facebook. As for devices, we found only two instances of wearables [117, 118] and two instances of ambient objects [66, 119]. In addition to the wearable electrode discussed above, the other example of a wearable device is in a paper by Sengupta et al. [118] that used a smartwatch to measure steps, in order to provide a feedback nudge to hospital patients to increase their mobility. The first example of an ambient object is from Hoffmann et al. with their connected dashboard in a truck [66]. Here, 104 truck-drivers received feedback on their driving style to increase their eco-efficient driving behavior [66]. The other example for this delivery channel is a connected bottle that overflows to give feedback to users who do not drink enough water [119].

2.3.2 Digital nudge context

To code the problem contexts of digital nudges, we focused on application domains on the one hand and desired outcomes on the other hand.

Application domains

To code application domains, we relied on categories identified in previous literature (i.e., [73, 74]) and adapted them to offer additional granularity. The following list of domains was established: privacy/security, e-commerce/marketing, social media, sustainability, recommender systems, crowdfunding, policy-making, education, innovation, work, education, health, and miscellaneous. The heat-map in Figure 2.5 shows the application domains in which each digital nudging pattern is studied.

The application domains in which digital nudges are used to influence choices or behaviors most frequently is privacy/security. Various papers in this context have focused on designing UIs to make privacy policies more accessible to users [52], facilitating user privacy interactions by modifying cookie consent notices [82], or nudging users into adopting electronic identification (eID) features [115]. However, e-commerce/marketing and various forms of sustainability behaviors are also common problem contexts.

Furthermore, we found very few examples of otherwise well-known patterns of digital nudges (e.g. friction, feedback, or warnings/graphics) being applied to any of the problem contexts in the bottom of Figure 2.5. A rare example of a study targeting the infrequent context of health came from the previously mentioned paper by Sengupta et al. where eight hospital patients were being evaluated over 91 days [118].

Outcomes

To code the outcomes of the studies, we used a more abstract level and looked at whether the objective was centered on changes in behavior or attitude. When studies investigated both of these higher-level outcomes or incorporated other levels such as physiological outcomes, we coded them as multi-level. The heat-map in Figure 2.6

Figure 2.5: Frequency of digital nudges per pattern and application domains (unit of analysis: nudges)

shows that digital nudges are most frequently used to change behavioral outcomes. An example of nudges targeting behavioral outcomes can be found in Oldenbeek et al. [120]. They showed how a simple feedback nudge in the form of a personalized

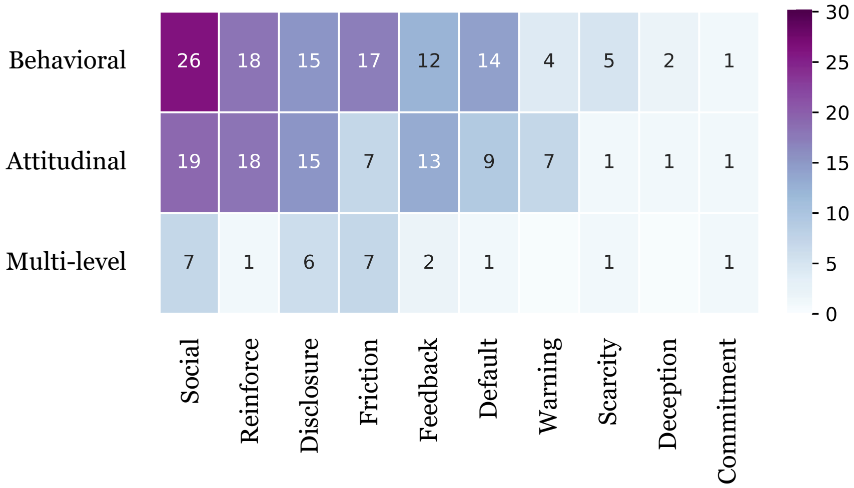

Figure 2.6: Frequency of digital nudges per pattern and type of outcomes (unit of analysis: nudges)

email on students' online lecture viewing progress had an impact on the behavior of the students. In short, this simple feedback made the students watch more of the online course content. An example for attitudinal outcomes can be found in Schöbel et al. [121]. With an online survey they studied Slack user preferences towards several digital nudging patterns (defaults, social, feedback etc.) in the context of privacy. By showing users a variety of nudging patterns and simply asking for their most vs. least preferred option they showed that users preferred to be nudged in the direction of privacy by defaults and red/green warning graphics. Multi-level outcomes are rarely examined in a single study. Note however that some papers employ several studies with different objectives that together gauge attitudinal and behavioral outcomes and these are labeled as multi-level in Table 2.1. Furthermore, since most nudges belong to studies with behavioral outcomes, we deepened our analysis by coding the type of behavior that was targeted with the nudge.

Using a simplified version of Fogg's behavior grid [122] we coded whether digital nudges that were focused on behavioral outcomes were used to investigate single or sustained occurrences of behavior, and whether the objective was to instigate a new behavior, increase/decrease a familiar behavior, or stop a behavior from occurring. As Figure 2.7 illustrates, most behavioral studies were focused on one-time outcomes aimed at increasing the intensity of a familiar behavior or on instigating a new behavior for the user.

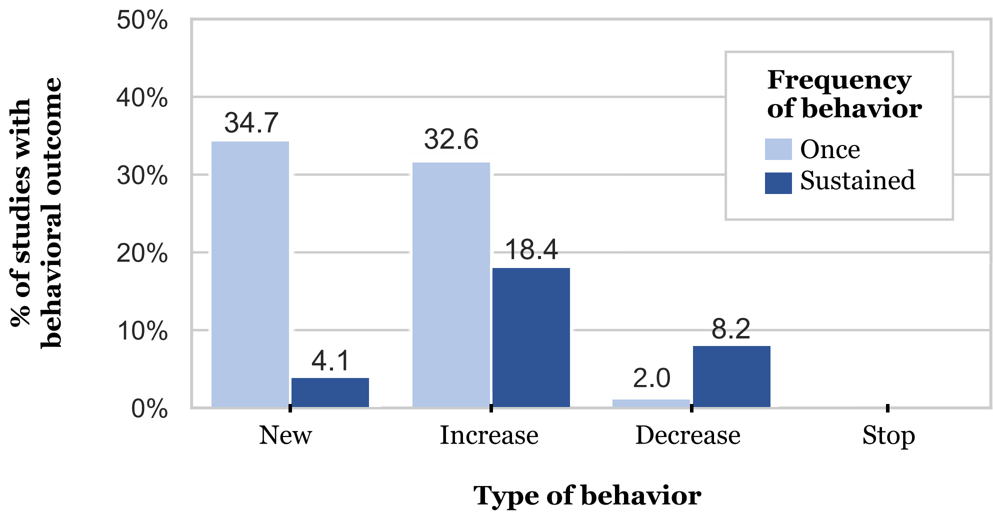

Figure 2.7: Percentage of behavioral outcome studies across targeted behavior types (unit of analysis: studies)

In other words, these studies measured the outcome of their digital nudge once. What is far rarer is studies where the outcomes of digital nudges are measured over a sustained period. The previously mentioned field experiment involving the driving behaviors of truck drivers was conducted over a 30-week period [66], while Kim et al. decreased app usage with their lockout task in a 3-week field experiment [65].

Another example was the combination of reinforcement and feedback nudges aimed at increasing the physical activity of hospital patients over a 3-month period [118].

2.3.3 Evaluation of digital nudges

To assess the types of empirical evaluations that were used in the literature, we coded the type of method and the type of sample, and we also assessed whether a different combinations of nudges were investigated.

Evaluation methods

We grouped studies according to their evaluation methods (e.g., online experiment, field experiment, online surveys, observational field study etc.). Figure 2.8 reveals that the most frequent methods for evaluating digital nudges were online experiments. Just as in the previous section, we still only found a few examples of digital nudges (e.g., feedback and warnings) being demonstrated or evaluated in field settings. Furthermore, qualitative evaluation methods are very rare in the domain of digital nudging. Some examples can be found in Kim et al., where they also conducted user interviews after a field experiment [65]. Another study elicited expert feedback [107], and Schilling et al. used both expert feedback and focus groups as evaluation methods [123].

Sample type

We investigated what type of sample was used for each evaluation (e.g. online panels, university students, partner-organizations). The heat-map in Figure 2.9 shows that when digital nudges were evaluated, a variety of sample types were used. Our analysis suggests that the use of online panels is now more common than reaching out to local university students. The most frequently used online panels in the literature were Mturk, Prolific, and Qualtrics. We also observe that several field studies or experiments are conducted with user data from a partner-organization (see for instance

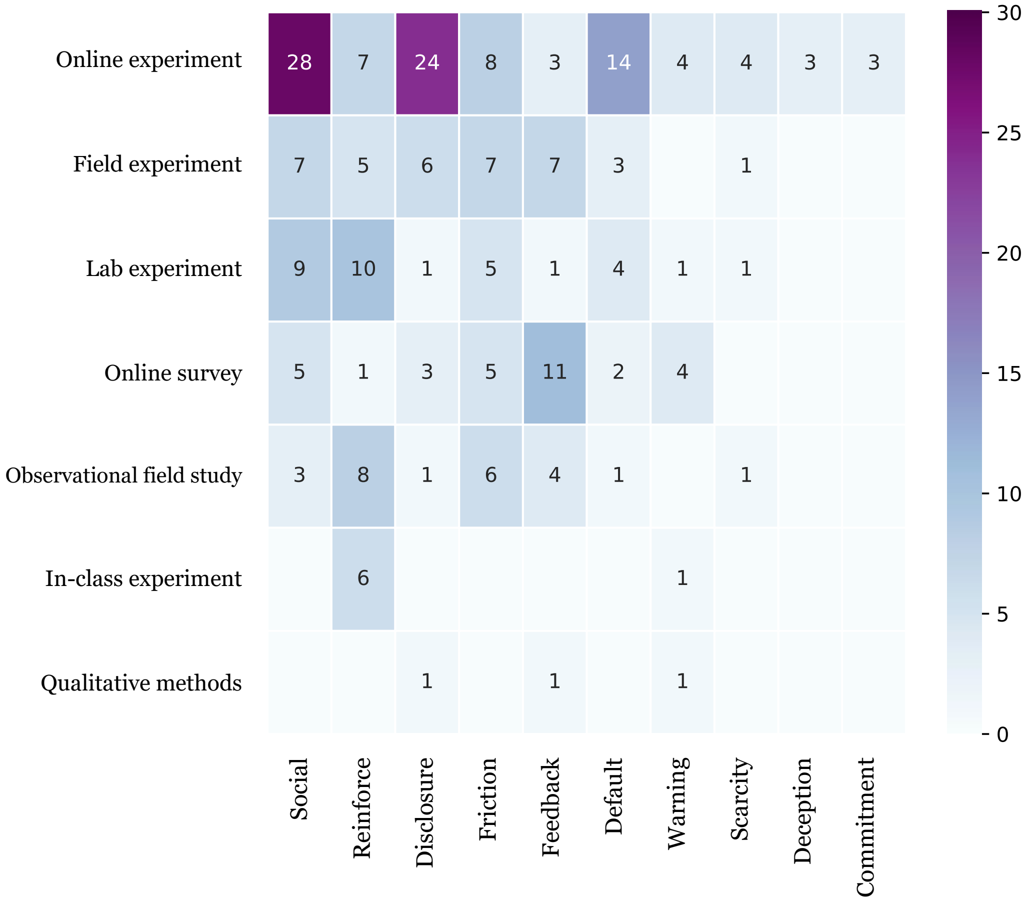

Figure 2.8: Frequency of digital nudges per pattern and evaluation method (unit of analysis: nudges)

[84], [52], and [111]). Figure 2.9 echoes the point that several patterns of nudges are not evaluated based on samples from various field settings. As for sample geography, around 69% of all studies sourced participants from either Europe or North America. Four studies were conducted in Asia, one study might have been conducted on a sample from South America [124], and no studies were explicitly sourced from the African continent. Roughly 27% of all studies did not explicitly disclose the geographical location from which their users were sampled. Evaluation methods and sample types also overlap in important ways.

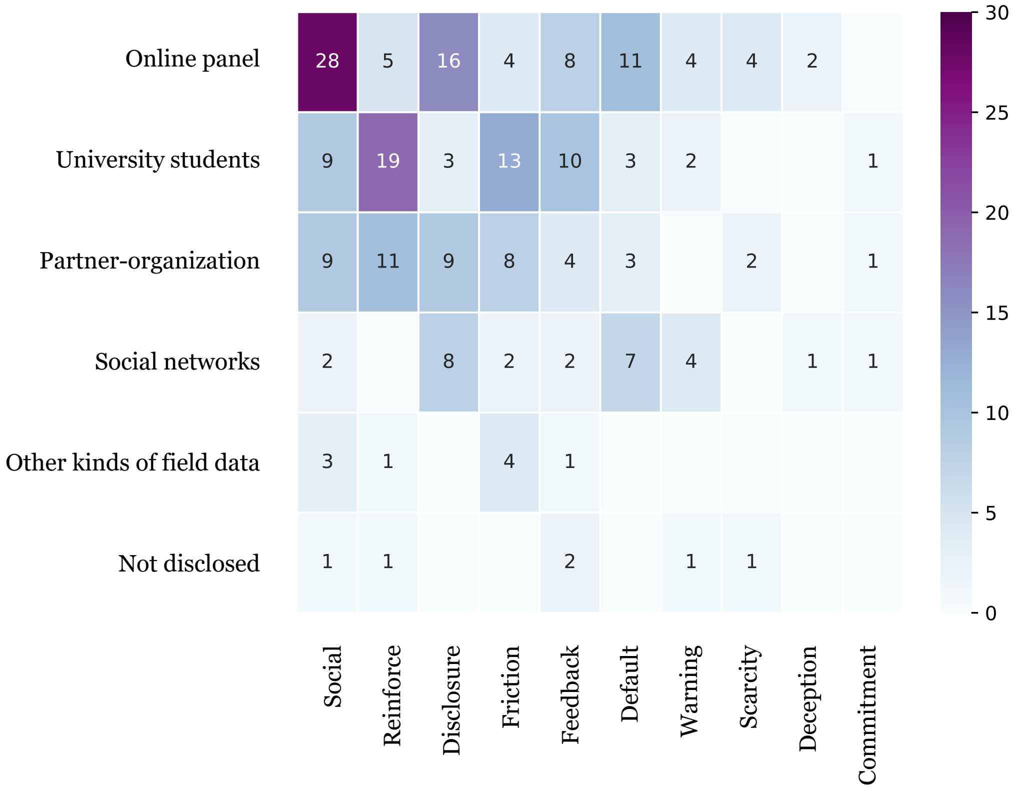

Figure 2.9: Frequency of digital nudges per pattern and sample type (unit of analysis: nudges)

For example, online experiments are often conducted with participants from an online panel, and field experiments often require a partner-organization.

Nudge combination evaluation

We assessed whether evaluations covered more than one nudge and whether their interaction effects were measured. The heat-map in Figure 2.10 highlights the number of times that digital nudge patterns were evaluated together in the same study. Depending on the design of such a study, this provides an opportunity to measure if the effect of one nudge pattern is dependent on the presence of another (i.e., statistical interaction). While a measurement of interacting nudge patterns depends on the evaluation method that is being employed it should be stated that when digital nudges are studied together, only in about a quarter of cases are statistical interactions measured and evaluated. The results in Figure 2.10 are obviously correlated to the number of times specific nudge patterns have been investigated in the literature. However, these results illuminate quite significant knowledge gaps around how different combina-

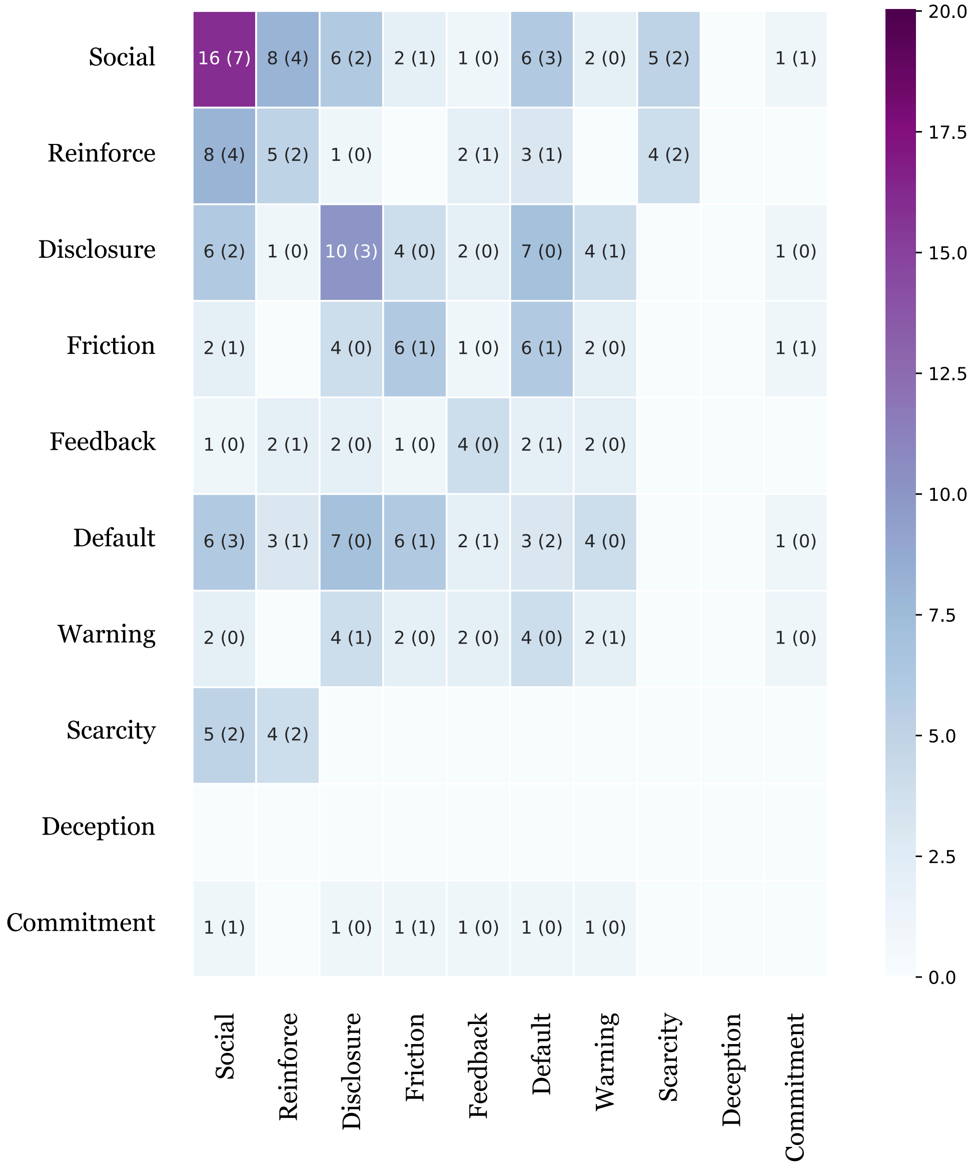

Figure 2.10: First number is the frequency of digital nudge patterns that were evaluated together, and the second number in brackets () is the number of times statistical interactions were measured (unit of analysis: studies)

tions of nudge patterns might impact user choices and behaviors. Good examples of studies where nudge patterns are used together and where statistical interactions are examined can be found in the previously mentioned papers (see for example [105, 88, 118]).

Summary of the digital nudging landscape

Table 2.1 provides an overview of the digital nudging landscape by showing the most relevant information that was coded from the selected papers in this review. The unit of analysis for this table is papers. For each paper, the table presents the nudge patterns that were investigated, the types of outcomes that were measured, the context, and the types of evaluation that were conducted. The table indicates whether the paper also investigated a partial or fully personalized and/or interconnected choice architecture. Further, it highlights the instances where a paper investigated particular delivery devices (i.e., wearable or ambient) or delivery channels (i.e., haptic). Finally, it indicates the number of studies conducted in each paper.

2.4 Discussion

This chapter presented a landscape of empirical research contributions on the growing topic of digital nudging through a systematic literature review. After an initial count of 638 papers related to "digital nudging", we conducted a thorough screening and eligibility process which led to the inclusion of 73 papers that were analyzed in detail along four dimensions: patterns of nudges, types of outcomes, types of contexts, and types of evaluations. Overall, the results of this review point to a highly heterogeneous research domain.

While the most common publication venues are associated with information systems and computer science, digital nudging is now being applied to a varied set of contexts, ranging from policy interventions [101] to the design of hydration-encouraging water bottles [119].

Strikingly, our results show that some common online nudges are hardly investigated. The most prominent example is scarcity, which is used by many service providers to nudge users to quickly purchase a service or product (e.g. "three seats left at this price" or "two rooms left for these dates"). Our results further show that field experiments are relatively rare, as they represent less than 16% of studies. This is again surprising given the ubiquity with which experimentation occurs online. In the industry, hundreds of experiments are being run daily to determine the efficacy of small visual changes in design features at various tech companies such as LinkedIn [125], Facebook [126], and Google with their infamous 50 shades of blue experiments, where the color of advertising links was live tested on users and reportedly led to USD 200 million a year worth of extra clicks [127]. This points to a possible gap between the digital nudges that are implemented in practice and the digital nudges that are actually researched. The imbalance or lack of research on certain nudges (e.g. deceive), types of behavior (e.g., decrease or stop), delivery channels (e.g. haptic), or device (e.g. wearable) and the lack of certain evaluation methods (field experiments) could also show a focus on what is convenient to measure rather than on what is relevant to measure (i.e., implemented in practice or forward looking). Prominent researchers have observed a similar phenomenon in the field of consumer behavior and asked to avoid a too narrow framing on most obvious or easy-to-measure behaviors (e.g., [128]). Below, we discuss our results structured along six key topics that emerged from our analysis. We also lay out future research questions that could be addressed to fill important knowledge gaps and open new research avenues.

2.4.1 Context and culture

Previous research on digital nudges focused on three problem contexts that take up almost half of the research effort (i.e., privacy/security, e-commerce/marketing, and social media). Research in the other contexts is much less prominent. The original conceptualization and applications of nudging by the authors of the book "Nudge" were deeply anchored in policy-making. It is thus surprising that only 7% of research on digital nudging is applied to this context. A similar observation can be made about the context of health. These disparities raise the following research question:

Future RQ1: Can the outcomes of a digital nudging be transposed to different classes of problems?

Geographic sample diversity is also lacking, as only five studies in the included papers have been conducted outside of Europe and North America. This should be a research priority if we want to understand whether the effects we study entail some forms of universalism, or whether digital nudges are bounded by cultural conditions. For instance, given that nudging has a specific political ideology (libertarian paternalism), could the effect of digital nudges be dependent on cultural aspects such as power distance beliefs (accepting and expecting power inequalities, [129, 130]) or cultural tightness (vs. looseness, [131]). More broadly, this raises the following future research question:

Future RQ2: Are the effects of digital nudges moderated by cultural contexts?

2.4.2 Choice architecture externalities

The results highlight that the majority of digital nudges have been researched in contexts that we call disconnected and non-personalized. This means that nudges are employed in situations where choice architectures are independent of one another and the choice architecture is not investigated in a context where personal data from the user is taken into account (i.e., non-personalized). While this increases effectiveness of evaluating the impact of various nudging patterns, it does not accurately reflect the digital sphere that users find themselves in every day and leaves open the following research question:

Future RQ3: How can the impact on users from interconnected and personalized choice architectures be better quantified and understood?

Indeed, Google's search results, Facebook's newsfeed, or video recommendations from YouTube are all highly personalized and interconnected choice environments and are likely capable of nudging users in one direction or another. In this context, several studies have been investigating how potential positive feedback loops leading to filter bubbles can be reduced [132]. On social media timelines, personalized nudges, which display relevant information more prominently, lead users to interact with it more. The increased interaction leads the algorithm to display more similar content and so on. In the end, a so-called filter bubble is created. In some cases, where the filter bubbles become problematic in terms of disinformation, social media platforms might use a warning nudge to break the user out of the vicious circle [132]. Another example, for the same purpose, is to present a feedback nudge to a user showing the balance of opinion on their timeline to raise awareness [133]. While this particular context has been under increased scrutiny over the past five years owing to the fake news phenomenon, research is still ongoing and calls for a better understanding on how platform design and individual choices combine to influence outcomes [134].

2.4.3 Personalization tools

Our review shows that personalized choice architectures are a promising research avenue, as researchers have started to investigate how different personal and psychological traits influence the effectiveness of nudges. However, the reviewed articles align with the observation of Caraban et al. [74] that personalized nudging remains understudied. While taking the digital privacy of users into account, this leaves open the following research question:

Future RQ4: How can effective personalized nudges be designed using minimal real-time contextual data?

The trade-off between large scale data collection to design highly personalized nudges and privacy-by-design one-size-fits-all nudging is an important aspect to consider. Technically, personalization can be achieved at a crude and transparent level by asking users about preferences. More advanced personalization can be achieved by (1) understanding which user characteristics (e.g., demographic data, psychological data) or contexts (e.g., location, current behavior, or time [135]) predict the efficacy of different nudges and (2) use adequate tools to identify when a user fits these characteristics or finds themself in a particular context. With digital technology, it becomes increasingly feasible to infer some of a user's behaviors online (e.g., is the user reading an article? [136]) and offline (e.g., is the user going home? [137]), but it remains a very difficult task, which requires creative solutions for many behaviors (e.g., researchers have suggested detecting whether a user is smoking by using acoustic properties of a smoking breath recorded through wearable devices [138]). Even contexts such as the current transportation mode of a user (e.g., car, bus, train, bike) are difficult problems, where effective solutions require creative combinations of different sensors (e.g., motion, sound, and vision [139]). Promising examples of such approaches have been successfully used in similar areas of research. For instance, in their work, Anagnostopoulou et al. [140] used a transportation mode detection system to personalize interventions to support sustainable behaviors of a user. Future research could further take advantage of novel devices and sensors to improve context and behaviour detection. This avenue of research by definition raises questions about privacy and data tracking. It also raises a broader ethical question. Nudging in general has been called into question and labeled manipulative [141].

Personalized digital nudging would go a step further, increasing their power to influence users to choose what the designer decided. Also, this is happening in a context where personal data is being proactively collected about users. Thus, even if digital nudges are, by definition, letting users have the final choice, one may argue on the one hand that artifacts that steer users in a specific direction with an ever-increasing probability (such as could be achieved by personalized digital nudges) constrain users' autonomy and even limit their freedom of choice. However, it could also be argued that if a nudge is in the interest of a user (as perceived by her or him), it would be unethical not to increase the probability of steering a user in a certain direction [142]. These considerations lead to the following broader question:

Future RQ5: What are the emergent ethical boundaries of personalized nudging?

2.4.4 Pervasive delivery

In our review we found only a very few instances of nudges using delivery systems other than visual interfaces on mobiles or desktops. This leaves open the following research question:

> **Future RQ6:** Are the effects of digital nudges moderated by the delivery device or channel?

Typical examples of such innovative delivery channels include wearable devices or ambient objects, and could offer interactions through visual, haptic, audio, or olfactory cues. At this stage, it is not clear to what extent these novel delivery channels could help design more effective nudges, even though there are some promising avenues for some of them. For instance, the good vibration study found that a haptic feedback nudge could reduce the time spent on social media by up to 20% in the short term [102]. If used well they can offer more subtle, but more visible and targeted, interactions than mobile or desktop interventions in certain contexts. Examples include an artistic digital screen to display electrical consumption directly on the kitchen counter [143], a teddy bear with a connected Raspberry Pi that tells its owner when to take their asthma medication [144], a virtual aquarium that visualizes user contributions to a knowledge management system in a shared office [145], or a connected smart mirror, which could not only serve as an ambient feedback object, but also include augmented reality features [146].

2.4.5 Symmetric effects

Only about half of the outcomes of the digital nudges related to actual behaviors (vs non-behaviors, which we broadly categorized as attitudes). This is a small proportion, given that behavior change is inherent to the original definition of a nudge. Also, almost 90 percent of behavioral studies were focused on instilling a new behavior or increasing an existing one. Only about ten percent focused on decreasing behaviors and not one on stopping behaviors. Interventions to decrease or stop behaviors are, however, common (for instance in public health). Future work could restore balance:

Future RQ7: Is a given digital nudge as effective at stopping or decreasing a behavior as it is at initiating or increasing a behavior?

In a similar vein, most research evaluated the effect of nudges on one-time behavior changes. However, the long-term effects of digital nudging are not yet well understood [74]. Simplicity can again partially explain this narrow focus on one-time behaviors: it requires more resources and more complex designs to monitor and analyze long-term behavior change, while it also poses several ethical hurdles to research. However, this longer-term perspective is important if we want to understand real phenomena, as behaviors online are often repeated over time. Also, in many contexts, users are exposed to the same nudge several times.

Future RQ8: Is a given digital nudge as effective at changing repeated behaviors as it is at changing one-time behaviors?

These two questions (i.e., RQ7 and RQ8) raise the issue of symmetry of effects of digital nudges on different outcomes. Current knowledge cannot provide clear answers given the imbalance of research output on the different types of outcome. Overall, future research could focus on investigating different types of behavior change as well as longer-term behavior change. This would more accurately reflect the phenomena actually occurring, despite providing a more challenging problem to solve for researchers.

2.4.6 Combined effects

As nudging gains traction in the private and public sphere, users become more likely to be exposed to several nudges at the same time [147]. However, our results show that the formal study of the interaction effect of digital nudges is not yet mature. Some nudging practitioners have attempted to evaluate how interventions interact, see for instance [148]. Also, a recent study showed that a combination of defaults and social influence nudges led to a stronger impact on compliance for the user than each nudge individually [105]. Yet, the study of the potential interaction effects due to the combination of different digital nudges is only at its infancy and still requires further empirical investigation. For instance, in the context of social media, are the effects of adding two types of frictions on the same social media platform simply cumulative or does one friction feature moderate the effect of the other? Also, more formal investigations of the underlying mechanisms related to potential interaction effects of nudges are yet to be made. This leaves open the following broader research avenue:

Future RQ9: How do digital nudges interact and through which mechanisms?

2.4.7 Limitations

Even though we followed the guidelines for systematic reviews carefully, the present chapter is not without limitations. As is inherent in a systematic literature review, its scope is limited to the search results of the initial query and the database selection. Performing backward searches (i.e. examining all the papers in the reference lists of the found articles) or forward searches (i.e. examining papers that referenced any of the found articles) could potentially augment the scope of the the current review. We recognize that there are other papers (e.g. [149, 70, 68]) that would be relevant to the digital landscape of nudging. Yet, the mentioned papers have not been included in this literature review as they never mention or reference digital nudging by name. We could have aimed for more exhaustivity by using search terms specific to nudge patterns whose authors did not call digital nudges (e.g. online commitment). However, for this approach, we would have needed to identify an exhaustive list of search terms, which is hard to attain and risked biasing our results by putting too much emphasis on one pattern over another. Our approach is more restrictive, limiting the results to research labeled as digital nudging, but this ensures exhaustivity within that label, which is our objective. Since the exact definition of what digital nudges entail remains relatively blurry, our research describes the landscape of what authors generally call digital nudging. Our hope is that this work may support future efforts that seek to further converge towards a unified definition of digital nudging, that would describe what digital nudging entails and what it does not, as the boundaries are not always consistent throughout the literature. Over the years, various definitions of nudging have grown more specific, yet more complex – see [60] for a discussion on this topic. Since the number of articles explicitly related to digital nudging had grown very steeply over the last five years, we judged that it was the right time to capture the evolution of digital nudging literature and map it for future researchers.

2.5 Conclusion

This chapter presented a systematic review of the literature on digital nudges. We provided a detailed analysis of 73 papers, based on 109 studies containing 231 digital nudges that had been demonstrated and evaluated by previous scholars. Our results show that while research on the topic has been very active, there seems to be a concentration of investigations in several specific areas of the landscape.

For instance, almost half of the research on digital nudging focuses on three contexts: privacy/security, e-commerce/marketing, and social media. Furthermore, several patterns of nudges seem to be hardly investigated, even though they appear to be widely used online, such as scarcity nudges (e.g., only one room left), deception nudges (e.g., a decoy option), or commitment nudges. The type of behavioral outcome is very rarely a decrease or a complete cessation of a behavior when nudges are applied in the literature. Moreover, while few single studies measure both attitudinal and behavioral outcomes, more papers conduct several studies to better triangulate outcomes.

Finally, while both interconnected and personalized choice architectures can differentiate digital from conventional nudges, most of the research is conducted in what we call disconnected and non-personalized choice environments. Based on these observations, we laid out nine future research questions to illuminate several unexplored patches in the digital nudging landscape that future researchers might want to explore.

Table 2.1: Digital nudge literature overview containing: Nudge pattern. Outcome. Context. Evaluation. Sample type. Personalization: partial (★), full (★★). Interconnected: partial (★), full (★★). Delivery channel other than visual through a mobile or computer screen (Delivery): Haptic channel (Haptic), Wearable device (Wearable), Ambient device (Ambient).

Paper	Nudge pattern	Outcome	Context	Evaluation	Sample type	Personalized	Interconnected	Delivery	Studies
[87]	Social, Disclosure	Attitudinal	Miscellaneous	Online Experiment	Online panel		★		1
[150]	Social	Attitudinal	E-commerce / Marketing	Online Experiment	Online panel	★	★		1
[151]	Reinforce	Attitudinal	Recommender Systems	Online Experiment	Online panel	★★			1
[152]	Reinforce, Scarcity	Behavioral	Crowdfunding	Online Experiment	Online panel	★	★		1
[153]	Social	Attitudinal	Social media	Online Experiment	Online panel	★	★		1
[154]	Social	Behavioral	E-commerce / Marketing	Online Experiment	Other field data		★★		1
[114]	Disclosure	Multi-level	Policy-making	Lab Experiment	Online panel		★★		1
[155]	Social, Disclosure, Feedback, Warning, Commitment, Default	Behavioral	Work	Online Experiment	Social networks				1
[119]	Feedback	Behavioral	Health	Field Study	University students			Ambient	2
[156]	Social, Friction, Default	Behavioral	Sustainability	Field Experiment	Social networks	★★			1
[52]	Friction	Multi-level	Privacy/Security	Online Experiment, Field Study	Online panel, Partner-organization				2
[157]	Friction	Attitudinal	Innovation	Lab Experiment	University students	★			1
[98]	Social, Reinforce	Multi-level	Education	Online Experiment, Field Experiment	University students, Online panel	★★	★★		2
[116]	Reinforce, Warning, Deception	Attitudinal	Privacy/Security	In-class Experiment, Online Experiment	University students, Social networks	★			6
[46]	Disclosure, Friction, Default	Behavioral	Crowdfunding	Field Study	Partner-organization				1
[158]	Reinforce, Social	Behavioral	Policy-making	Field Study	Partner-organization		★		1
[159]	Disclosure	Attitudinal	E-commerce / Marketing	Online Experiment	University students				1
[97]	Reinforce	Attitudinal	E-commerce / Marketing	Lab Experiment	University students		★		7
[160]	Feedback	Attitudinal	Privacy/Security	Online Survey	Online panel	★			1
[94]	Social	Behavioral	Health	Online Experiment	Online panel	★			1
[101]	Friction	Behavioral	Policy-making	Field Study	Other field data		★		1
[161]	Feedback	Multi-level	Recommender Systems	Field Study, Online Survey	University students				4
[162]	Social, Reinforce	Multi-level	Recommender Systems	Field Experiment, Online Survey	Partner-organization				3
[100]	Disclosure	Behavioral	Social media	Online Experiment	Social networks				1
[163]	Reinforce, Default	Behavioral	Sustainability	Lab Experiment	University students	★			1
[66]	Feedback	Behavioral	Sustainability	Field Experiment	Partner-organization	★		Ambient	1
[164]	Disclosure	Behavioral	Social media	Field Experiment	Partner-organization				1
[165]	Disclosure	Behavioral	Sustainability	Online Experiment	Social networks				1
[117]	Feedback	Behavioral	Miscellaneous	Field Experiment	Not disclosed			Haptic, Wearable	2
[105]	Social, Default	Behavioral	E-commerce / Marketing	Online Experiment	Online panel	★	★		2
[106]	Default, Warning	Behavioral	E-commerce / Marketing	Lab Experiment	Online panel	★★			1
[166]	Social, Disclosure, Default, Warning	Attitudinal	Privacy/Security	Online Experiment	Online panel				1
[167]	Disclosure	Attitudinal	Sustainability	Online Experiment	University students	★★			1
[168]	Social, Reinforce, Scarcity	Behavioral	E-commerce / Marketing	Online Experiment	Not disclosed				1
[169]	Social	Multi-level	Social media	Online Experiment	Online panel		★		2
[65]	Friction	Multi-level	Miscellaneous	Field Experiment, Qual. methods	University students	★★			2
[88]	Social, Disclosure	Multi-level	Social media	Online Experiment	Online panel	★	★		2
[86]	Social	Multi-level	Recommender Systems	Lab Experiment	University students	★	★		1
[170]	Friction, Warning	Attitudinal	Privacy/Security	Online Survey	Social networks	★	★		1
[113]	Social, Commitment	Attitudinal	Sustainability	Online Experiment	University students		★		1
[171]	Feedback	Behavioral	E-commerce / Marketing	Lab Experiment	University students	★★			1
[172]	Reinforce	Behavioral	Crowdfunding	Field Study	Partner-organization				1
[173]	Friction, Default	Behavioral	E-commerce / Marketing	Lab Experiment	University students				1
[174]	Social, Disclosure, Reinforce	Attitudinal	Sustainability	Online Experiment	University students				1
[175]	Friction, Default	Behavioral	E-commerce / Marketing	Online Experiment	Online panel	★			1
[176]	Disclosure, Default	Multi-level	Crowdfunding	Online Experiment, Field Study	Online panel, Other field data	★			3
[102]	Feedback	Behavioral	Social media	Online Experiment	Online panel	★★		Haptic	1
[124]	Default	Attitudinal	Policy-making	Online Experiment	Social networks				1
[177]	Reinforce	Behavioral	Education	In-class Experiment	University students				1
[178]	Friction	Behavioral	Miscellaneous	Online Experiment	Online panel	★			1
[85]	Friction, Feedback, Warning	Attitudinal	Health	Online Survey	University students	★			1
[179]	Friction	Attitudinal	Social media	Online Survey	University students, Partner-organization		★		2
[180]	Feedback	Behavioral	Education	Field Study	University students	★★			1
[181]	Social, Friction	Multi-level	Innovation	Online Experiment	University students	★	★		1
[123]	Disclosure, Feedback	Attitudinal	Work	Qualit. methods	Partner-organization				2
[182]	Disclosure	Behavioral	Privacy/Security	Online Experiment	Online panel	★			1
[110]	Scarcity, Deception	Behavioral	Crowdfunding	Online Experiment	Online panel				3
[115]	Social, Default	Behavioral	Privacy/Security	Online Experiment	Online panel	★	★		1
[121]	Social, Disclosure, Feedback, Warning, Default	Attitudinal	Privacy/Security	Online Survey	Social networks				1
[183]	Friction	Behavioral	Privacy/Security	Online Experiment	University students	★			1
[184]	Disclosure	Attitudinal	Sustainability	Online Experiment	Online panel	★★			1
[118]	Reinforce, Feedback	Behavioral	Health	Field Experiment	Other field data	★★		Wearable	1
[185]	Reinforce, Default	Attitudinal	Innovation	Lab Experiment	University students				1
[186]	Reinforce, Feedback, Default	Behavioral	Innovation	Field Experiment	Online panel	★			1
[84]	Social, Scarcity	Multi-level	Social media	Field Experiment, Lab Experiment	Online panel, Partner-organization		★★		3
[187]	Default	Attitudinal	Sustainability	Online Experiment	Online panel	★			1
[188]	Friction, Commitment	Multi-level	Miscellaneous	Online Experiment	Partner-organization	★			1
[83]	Feedback	Multi-level	Work	Field Experiment	Partner-organization	★★			1
[82]	Disclosure, Friction, Default	Multi-level	Privacy/security	Field Experiment, Online Survey	Partner-organization				4
[120]	Feedback	Behavioral	Education	Online Experiment	University students	★★	★		1
[107]	Warning	Attitudinal	Social media	Qual. methods	Not disclosed	★★	★		1
[111]	Social, Reinforce, Scarcity	Multi-level	Crowdfunding	Online Experiment, Field Study	Online panel, Partner-organization	★	★		2
[189]	Friction	Multi-level	Innovation	Lab Experiment	University students	★			1

3 Nudging for Online Privacy Awareness and Informed Consent

A version of this chapter has been published as: **Kristoffer Bergram**, Valéry Bezençon, Paul Maingot, Tony Gjerlufsen & Adrian Holzer. *Digital Nudges for Privacy Awareness: From consent to informed consent?.* 2020. In Proceedings of the European Conference on Information Systems (ECIS '20). AIS, Marrakech, Marocco.

Contents

3.1 Introduction

In this chapter we will explore what we previously referred to as a disconnected and non-personalized digital choice architecture. When browsing online, users often find themselves in a ubiquitous choice environment right before they are about to use any kind of digital service. With current regulations, online companies are forced to require informed consent from users before they can harvest their data in exchange for their 'free' services. This informed consent usually involves some complex agreement in the ToS and PP regarding when, how and why personal data is being collected and shared. Internet users can view these agreements if they desire before clicking the 'I agree' button. In reality, only a small number of people ever click the adjacent link to the ToS and PP. Most people just proceed, and thereby perpetrate what is known as the biggest lie on the Internet: I have read and agree to the terms and conditions [190].

3.1.1 Poor ability and motivation

While fruitful research has been conducted related to macro-level privacy concerns [191, 192], we will leave privacy calculus models aside and focus more on the observed behavior of users in this chapter. When it comes to explaining why Internet users tend to not read the ToS and PP of digital services, the scientific jury is still out. Several factors might explain users' dwindling ability and motivation to read the ToS and PP. Previous research on this topic highlights information overload on the part of the user [190, 193], that most users have "nothing to hide" [190], and the fact that these complex agreements are very hard for the average Internet user to understand [190, 193, 194, 195]. Even while leaving out the above factors, online privacy seems to be one of the contexts where many of the cognitive dials are turned in the wrong direction for optimal human decision-making [196].

Instead of having a rational wind at their back when making decisions about their online privacy – users seem to have a torrent of biases working against their better judgement (e.g. [196, 197, 198, 199, 200, 201]).

3.1.2 Information and power asymmetry

To further compound the confusion of Internet users, there is a large information gap between the providers and the end-users of digital services [197]. That is, service providers know with a high degree of specificity what data they want to collect from the user, whereas the user tends to have little to no idea about what data they share with the provider or other third parties [202]. Users also tend to keep better track of the benefits of an online service rather than the privacy risks that might be associated with that service [203]. Scholars have also suggested that Internet users have trouble distinguishing between their own publication control and the control related to the access of their personal information [198].

To illustrate: If one decides to publish an album of their kids' and share this with their Facebook friends – this does not necessarily mean that only their friends will have access to this data and its associated metadata. However, this distinction is hard to make for many Internet users. On top of this, earlier findings suggest rather deep misconceptions among Internet users. One survey conducted in the US suggested that 62% of the sample thought that if a website had a PP, it meant that the site could not share one's data with other companies [204]. On a general level, Internet users also seem to have a poor understanding of how their personal data is connected to the economics of these free and data-driven online services [205].

3.1.3 Towards better choice architecture

Our decisions are often influenced by the choice architecture that we happen to be presented with [206]. The insight that choice environments can affect the likelihood of certain decisions and their associated behaviors have given rise to the concept of "nudging" [75]. In the context of HCI, digital nudging refers to the use of UI design features that guide people's choices or behaviors in online decision environments [27]. Online privacy awareness is a promising problem for software designers to tackle, especially from the standpoint of choice architecture and nudging [198]. Several studies have already been conducted in the context of digital nudging and online privacy [207, 202, 208, 170].

This chapter will focus exclusively on the choice architecture that potential users encounter right before they decide to join or use an online service. This is the first chance users have to inform themselves about what privacy guarantees that are offered (or lack thereof).

Yet, research suggests that most users simply ignore this opportunity and thereby perpetrate the biggest lie on the Internet [190]. In this chapter, we address this problem by employing design considerations related to digital nudging. Specifically, we will investigate how system and software designers can improve user's privacy awareness through design modifications to the current choice architecture that is being used by leading social media brands such as Facebook [209], LinkedIn [210] and Twitter [211]. This leads to us to our overarching research question:

RQ: How can digital nudging improve the privacy awareness of users?

3.2 Methodology

To investigate this topic, we follow a design science research methodology (DSRM) in accordance with the steps outlined by Peffers et al. [212]. The rest of the chapter is structured as follows: Section 3.3 discusses the problem statement and Subsection 4 defines the objectives of a solution in the context of previous literature. Section 3.4 presents the design of a novel choice architecture for raising privacy awareness. We then present an experimental evaluation of the previously presented designs in section 3.5. Section 3.6 shows a real demonstrator of one of the novel choice architectures with the example of the IKEA Place app and discusses the observed user behavior in the field. Finally, the last two sections feature a discussion and concluding remarks for the research community.

3.3 Backgound

At the turn of the 21^{st} century, countries around the globe built the legal infrastructure to accommodate online privacy, electronic contracting and consumer protection [213]. During the last two decades, the means of extracting, cleaning, warehousing, analyzing and monetizing consumer data have arguably evolved [214]. Yet, the choice architecture that has been provided for consumers to 'agree' or to 'not agree' to these practices seems rather fixed.

3.3.1 Shrinkwrap, clickwrap, browsewrap and quick-join

Kunkel [213] presented case law distinctions between *shrinkwrap*, *clickwrap* and *browsewrap* that still provide an instructive background today. These distinctions can be thought of as different historical choice architectures to what eventually became the de facto interface for online privacy. Kunkel [213] outlined that before the Internet matured, shrinkwrap licenses got their name from the clear plastic wrapping that enclosed software packages. The packaged software contained a notice that by tearing open the shrinkwrap, the user agreed to the software terms enclosed within. The term clickwrap agreements emerged when software vendors began distributing software by means other than CD's as when the software was downloaded over the Internet. During installation or first use of an application, the classical dialog box containing the terms of the license would open for the user to read. The user was then asked to signal their informed consent by clicking 'I agree' or 'I do not agree' [213]. The last category called browsewrap agreements are not as overt as the former. They do not appear on the screen and the user is not compelled to accept or reject the terms in order to proceed with the installation, download or sign-up procedure. A browsewrap agreement only appears as a link that is accessed by clicking i.e., it is optional and not required to view the actual agreement. In the early days of the Internet, this choice architecture or form of contract was successfully challenged in US courts due to the legal reason that it lacked a reasonable notice and that consumers required a more earnest opportunity to review the terms they agreed to [215].

Today, several of the largest online service providers (e.g. Facebook, Twitter, Microsoft) offer a fusion of the browsewrap and the clickwrap architecture. One has to indicate their agreement with the click of a button but viewing the actual contents of the ToS and PP can still be a matter of choice, meaning that the agreements are at least readily available through an adjacent link. Obar and Oeldorf-Hirsch [216] now refer to this fusion category as a "quick-join" option while noting how services like Instagram and Twitter employ this choice architecture to ensure that prospective users can join quickly without having to be bothered by scrolling through any terms or policies before accepting them.

While the current quick-join environment speeds up the collection and sharing of personal data, it does little to facilitate the informed choices of Internet users. With this choice architecture, users are given a timely prompt in the form of a link right before they agree to any potential terms or privacy policies. However, recent research highlights that the majority of users rarely act on this signal.

3.3.2 The problem and reality of online privacy

A few data points from Obar and Oeldorf-Hirsch [190] will suffice to underline the main problem that this chapter is addressing. In the process of joining a fictitious social networking service, 74% of a sample of US college students joined without even looking at the privacy policy. Of those that did read it, the average reading time was 1 minute and 14 seconds. Obar and Oeldorf-Hirsch [190] estimated that it should have taken roughly 30 minutes to read the full PP that was used in their study. An even more disturbing finding was that less than 5% of the sampled participants expressed concerns around the fact that the ToS of the social networking service outlined that their data would be shared with prospective employers, insurance companies and the National Security Agency (NSA). Additionally, according to the ToS, each participant agreed to sign over their first-born child as intellectual property to the social networking service. Fortunately, the enslavement of an unborn child would not be upheld in a modern court system. However, complex data sharing practices are a real ubiquitous concern when it comes to online privacy.

As mentioned in the introduction, there are several explanations as to why we tend not to read the ToS and PP of online service providers. Research on this topic suggests that the primary factors explaining our lacking motivation and ability to look at online ToS / PP relates to information overload [190, 193], that most of us have "nothing to hide" [190], and the sheer difficulty of reading and understanding their contents [190, 193, 194, 195]. The reasons why we tend to not read online terms and policies can also be illuminated through a more psychological lens. Studies in economic and behavioral science suggest that we often place less weight on the future than on the present, this systematic bias is known as hyperbolic discounting [217, 218]. In short, when keeping other factors constant, our preferences related to pay-offs and hazards tend to change depending on how close these pay-offs and hazards are in terms of time. Internet users seem to stay true to this inconsistency when they make decisions related to their online privacy. Scholars have argued that people also put less weight on long-term risks and losses while acting in privacy-sensitive situations [219].

This fact makes people likely to trade their long-term privacy for relatively modest short-term pay-offs [220]. Another study highlighted how a nationwide sample of people from Singapore displayed a strong optimistic bias regarding their online privacy [200]. The participants in the sample tended to judge themselves less vulnerable than "others" to online privacy infringement. To summarize, the current choice architecture around online privacy already offers a small prompt in the form of an adjacent link that one can click to review the ToS and PP before using an online service. Previous research suggests that this privacy dialog box rarely engages Internet users to inform themselves before they consent to these terms and policies. Our contention is that by employing design considerations related to nudging and persuasive technology, this problem can hopefully be mitigated.

3.3.3 Digital nudges and persuasive technology

There is still a lot of debate around what exactly qualifies as a nudge [63, 141, 58]. However, Thaler and Sunstein [206, p. 6] defined a nudge as "... any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives". If a software designer works with any aspect of HCI, they are responsible for organizing the digital context in which their users make decisions [6]. That is, they are designing choice architecture.

The concept of 'digital nudging' simply refers to when peoples' choices or behaviors are being guided by UI design features in digital decision environments [27]. The use of digital nudges has been studied in a range of online contexts such as reward-based crowdfunding [110], social sharing [164], privacy [170] and digital addiction [221]. Digital nudging is becoming increasingly relevant because people are making more and more consequential choices in online decision environments.

The study of computers as persuasive technologies or 'captology' is an inquiry into the overlap between computer science and the psychological theories related to persuasion [222]. One of the important outputs of that research stream is Fogg's Behavioral Model (FBM). As a model, it offers three wide paths to aim for when persuasively designing for a new target behavior: Motivation, ability and prompts. That is, FBM emphasizes that to perform any target behavior, a person must be sufficiently motivated, have the ability to perform the behavior and be appropriately prompted or reminded to perform the behavior [223]. In other words, FBM is predicated on introducing pre-designed prompts while taking the user's ability and motivation into account. These pre-designed prompts come in three different varieties: *Signals* that remind the user of the task at hand, *sparks* that aim to increase the user's motivation to perform the task and *facilitators* that simplify the task for the users [223]. Other scholars have already attempted to combine FBM with the notion nudges. Caraban et al. [74] reviewed and categorized 23 distinct mechanisms of nudging and presented them in a framework for technology-mediated nudging. They also grouped these 23 nudging mechanisms into the three types of prompts proposed by FBM. Relevant to the current research, Caraban et al. [74] labeled one nudge as *throttling mindless activity* and categorized it as a spark in the context of FBM. That is, this nudge can be used to modify users' motivation in tasks were users tend to mindlessly proceed. Further, they labeled another nudge as *suggesting user alternatives* and categorized it as a facilitator to increase users' ability i.e., providing an attractive alternative to users can nudge them away from making a bad choice. These two nudges will be used to facilitate online privacy awareness.

3.4 Designing for Informed Consent

To assess how software designers can improve users' online privacy awareness the latter concept needs to be further defined. To increase users' online privacy awareness, users need to gain some meaningful facts about how their personal data is being processed. Such facts are addressed within the ToS and PP of online services. To increase their online privacy awareness users therefore need to view, read and recall some of the facts from the ToS and PP. This is also what we mean by moving users towards more informed consent. The quick-join environment offers a good benchmark and is consistent with several major online service providers like Facebook or Twitter [216]. Yet, only a minority of users respond to this design (by actually clicking the adjacent link to the ToS and PP).

Through the lens of FBM, what seems to be lacking is a privacy dialog box that sparks the users' motivation and increases the users' ability to inform themselves about their online privacy. Our design solution aims to take these two aspects into account by introducing two nudges. First, we suggest to spark users' motivation by throttling mindless activity. The current quick-join environment offers a small signal in the form of a link right before users agree to any potential terms or privacy policies. This is a type of choice architecture that resembles an opt-in policy where ignorance is the default, see Figure 3.1 A) on the next page. The quick-join choice architecture asks the users for consent but does not confront users with the fact that they have actively chosen to ignore the ToS and PP by not clicking the link. They have to opt-in to inform themselves. Prior research in the context of organ donations has demonstrated that a neutral choice compared to an opt-in policy is far more effective in soliciting organ donations [224]. A problem framed as an explicit choice between two equal alternatives compared to an opt-in policy is a way to increase the salience of the signal that is associated with the target behavior [225]. This can be achieved by designing a privacy dialog box with two buttons instead of an adjacent link. Such a choice architecture can be represented as a fork in the road, see Figure 3.1 B). This leads to the following hypotheses:

$H1_a$ A privacy dialog box that prompts users to make an explicit choice will increase the proportion of users that view the ToS and PP compared to a control group with an opt-in choice architecture.

$H2_a$ A privacy dialog box that prompts users to make an explicit choice will increase reading time of ToS and PP compared to a control group with an opt-in choice architecture.

$H3_a$ A privacy dialog box that prompts users to make an explicit choice will increase privacy policy recall compared to a control group with an opt-in choice architecture.

Secondly, to facilitate the users' ability to inform themselves about their own online privacy, we can suggest more alternatives related to how they view a ToS and PP. Solutions such as making content more intelligible and reducing information overload can often be facilitated by system and software designers. The design consideration of suggesting alternatives has been successfully employed in diverse contexts such as increasing the security of user's passwords [226], facilitating college choices [227] and increasing healthy food choices [228]. In the context of online privacy this can be accomplished by adding a brief summary of the full terms and privacy policy. Additionally, this kind of alternative has the added benefit of being a middle-option which increases its likelihood of being chosen [229, 230]. This choice architecture can be illustrated as a roundabout, see Figure 3.1 C). This leads to the following hypotheses:

$H1_b$ A privacy dialog box that prompts users to make an explicit choice and offers a summarized alternative will increase the proportion of users that view the ToS and PP compared to a control group which only prompts users to make an explicit choice.

$H2_b$ A privacy dialog box that prompts users to make an explicit choice and offers a summarized alternative will increase reading time of ToS and PP compared to a control group which only prompts users to make an explicit choice.

$H3_b$ A privacy dialog box that prompts users to make an explicit choice and offers a summarized alternative will increase privacy policy recall compared to a control group which only prompts users to make an explicit choice.

The two proposed nudges can now be illustrated within the context FBM as in Figure 3.1 D). The downward slope in the model is what Fogg [231] refers to as the action line. FBM highlights that prompts or nudges that occur on the right side of this line are more likely to achieve a specific target behavior. The standard quick-join choice architecture occurs mostly on the left side of the action line i.e., only moving

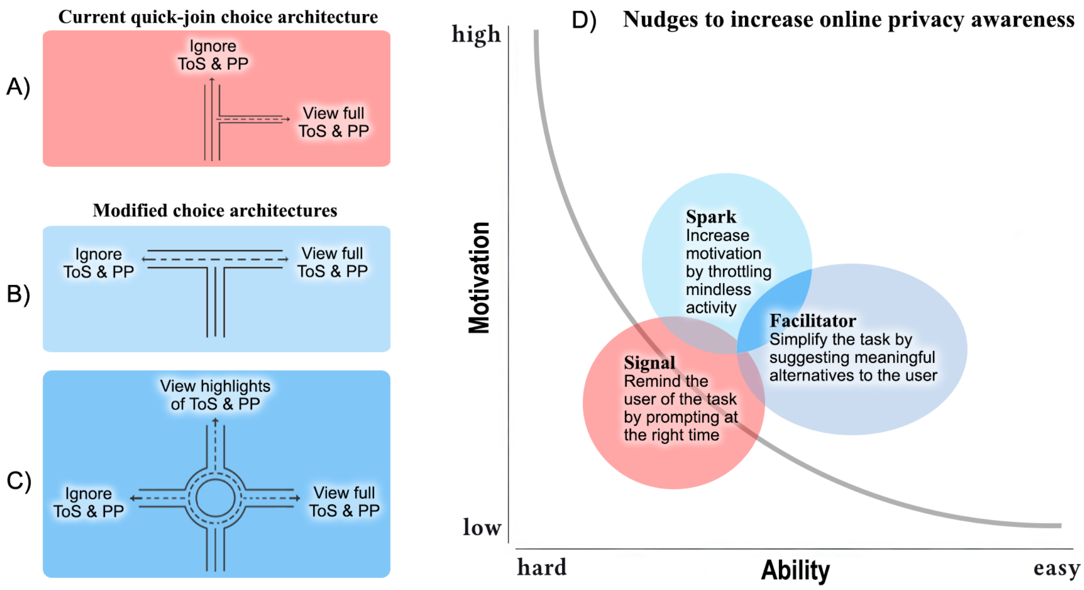

Figure 3.1: The choice architectures and the nudges of the proposed design solution mapped into FBM.

a minority to the right side of the line. Our argument is that by designing nudges that also target ability and motivation, we can increase the online privacy awareness of most users. The next subsection outlines how these nudges were employed in the final design solution.

3.4.1 Proposed user interface

Figure 3.2 presents a privacy dialog box (PDB) that implements both of the previously described nudges. The middle image shows a prompt that uses the roundabout choice architecture, i.e., providing an explicit choice while adding a summarized alternative of the ToS and PP. The 'Nothing really' button is the same choice as clicking 'Next'. The 'Just the highlights' button leads to a summarized version of the ToS and PP. The 'Everything' button features the full content of the ToS and PP. The proposed choice architectures will be described in more detail in the following subsection. We will use the quick-join design as a baseline and compare it on a number of dependent measures to two modified digital choice architectures.

There are a couple reasons why we use the quick-join environment in this study: First, for the control group this is currently in keeping with several major online service providers (cf. Facebook [209], Twitter [211], LinkedIn [210]). Second, in the modified choice architectures the aim is to change the target behavior without forcing users to comply to a given behavior i.e. viewing, reading or remembering the ToS and PP. In other words, users are free to choose their behavior and the modified choice architectures are aimed at nudging them towards a more informed type of consent.

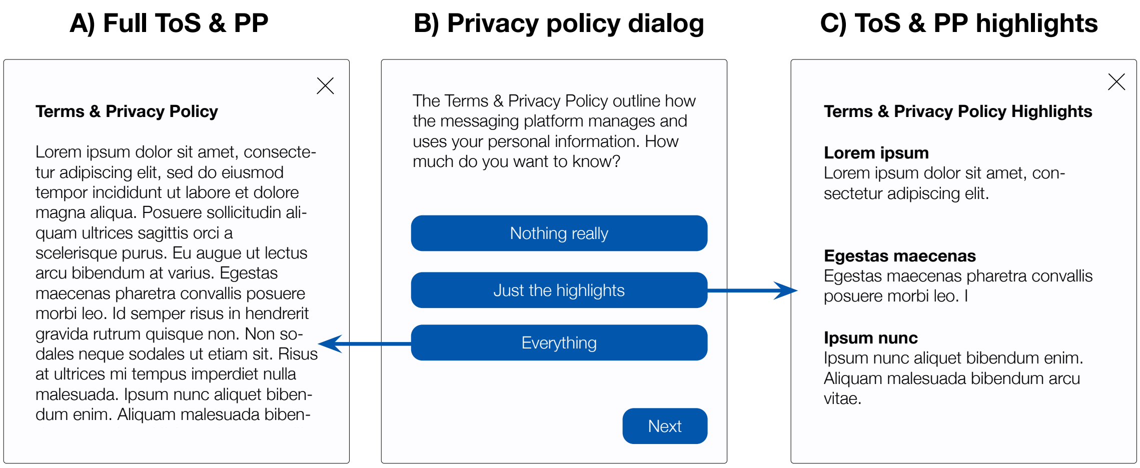

Figure 3.2: Mockup describing the proposed digital nudges in the privacy dialog box.

3.5 Evaluation

To investigate how software designers can improve user's privacy awareness we used the quick-join standard as a benchmark and compared it on a number of dependent measures to two modified choice architectures through an experimental design.

3.5.1 Procedure

The data collected for the experiment was drawn using Amazon's Mechanical Turk (Mturk) in accordance with Paolacci et al. [232] recommendations. Participants were filtered based on the location provided by their IP addresses, and only selecting participants from the US Mturk population. From Mturk, the online participants were directed to an anonymous Qualtrics link. The participants were asked to take a survey and perform a usability evaluation of a third-party online messaging platform. The online participants were not made aware that the research concerned nudging or online privacy. During the data collection, a double-blind procedure was also used, i.e., the researchers were unaware of which group each participant ended up in and this knowledge was also hidden from each participant.

The online participants were selected from the US MTurk population using the following qualifications: HIT approval rate $> 98\%$, number of HITs approved > 5000. A pilot experiment ($n = 56$) consisting of Swiss students suggested that the experiment took a maximum of 15 minutes to complete. Based on those results, the Mturkers were compensated \$2 US for their participation. The online participants first had to read the instructions and agree to the conditions of the study. After this, each participant was confronted with a randomly assigned choice architecture, see Figure 3.3. These designs will now be referred to as privacy dialog boxes (PDB). Apart from these designs, everything else in each experimental group was identical. Immediately after the participants had interacted with the PDBs they received 10 questions regarding the contents of the ToS and PP. After that, they performed a usability evaluation of a messaging platform and answered some more questions related to their general online privacy concerns, the trust they felt towards the messaging service and demographical information. At the end of the study, all of the participants were debriefed.

3.5.2 Measures

To test our hypotheses the following independent measure was devised: *choice architectures*. The following dependent measures were also used: *Viewed ToS and PP, Reading Time* and *Recall*. These measures will be operationalized below. Two control measures were also used: *General concern for privacy* from Schumann et al. [233] and a question regarding the level of *trust* that users felt towards the messaging service measured on a Likert-scale from lowest (0) to highest (10).

Digital Choice Architectures

Figure 3.3 below presents the three experimental conditions. PDB 1 (control) is an example of the current quick-join choice architecture where users have to opt-in to see the the ToS and PP like in Figure 3.1 A). PDB2 was nudged by an explicit choice i.e., a fork in the road as in Figure 3.1 B). PDB3 was also nudged with an explicit choice but with an additional suggested alternative i.e., a roundabout design shown in Figure 3.1 C).

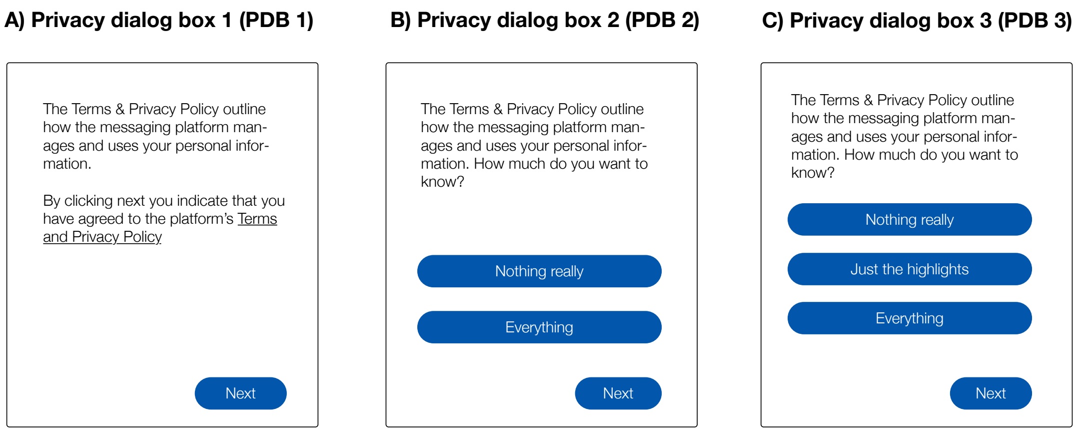

Figure 3.3: The digital choice architecture and design for each of the experimental groups.

Viewed ToS and PP

As Figure 3.3 illustrated, when the participants arrived at one of the randomized PDBs, the ToS and PP were hidden. The users had to click either a link or a button to view them. This dependent variable measured that click as a proxy of whether the ToS and PP was viewed. This created a binary outcome (viewed vs. did not view). As the results section will later show, this variable can be dissected further by separating the proportion of participants who clicked on the 'Highlights' button in PDB3, see Figure 3.3 C) above. One could also just click the 'Next' button on each dialog box to ignore the terms and policies. In other words, the choice to completely ignore the ToS and PP was not constrained by any of these designs.

Reading Time

This variable measured the time that each participant spent on the privacy dialog boxes. The count started when they arrived and stopped once they clicked the 'Next' button shown in Figure 3.3. This variable was measured in seconds and was considered a proxy of reading time.

Recall

To measure recall of the ToS and PP contents, we devised a battery of 10 multiple choice questions with 5 alternatives each. Each question concerned a specific detail related to users' personal data and all of these facts could be found within both the full and summarized versions of the ToS and PP. The right answer gave one point and the wrong answer gave zero points creating a scale from 0 – 10 that could readily be transformed to a percentage of correct answers.

3.5.3 Results

In total, 186 online participants were gathered for the online experiment. Three of these participants were excluded from the final analysis because they failed to submit a unique code that was generated for each run of the experiment. Out of the remaining sample ($n = 183$) there were 84 females, 98 males and one participant specified “other” as their gender. All the participants were between 19 and 68 years old ($M = 35.2$, $SD = 8.8$). The reporting of these results conform to the statistical standard of significance described by Benjamin et al. [234]. Since the dependent variables of this study did not conform to an approximate normal distribution, all inferential conclusions will stem from non-parametric tests. For the two control measures related to general privacy concerns and trust, no significant differences were found between the three choice architectures. Due to the different proportions of participants between the control group and the two experimental groups, a χ^2 test for the goodness-of-fit was conducted. The test indicated that there were no significant differences in the proportion of participants in the three randomized experimental groups $(50, 67, 66)$ as compared to the expected proportions of $(61, 61, 61)$, χ^2 $(df = 2, n = 183) = 2.98, p = .24$.

How choice architecture affected views (H1$_a$ & H1$_b$)

In relation to the above research hypotheses, a χ^2–test of independence highlighted a significant association between the three choice architectures and whether the participants clicked to view the terms and privacy policy, χ^2 $(df = 2, n = 183) = 27.27, p = 0.000$, Cramer’s V = 0.386.

Follow-up tests with a Bonferroni correction revealed a significant difference in the proportion of clicks between PDB1 and PDB2, χ^2 $(df = 1, n = 117) = 7.88, p = 0.005$ and a suggestive difference between PDB2 and PDB3, χ^2 $(df = 1, n = 133)$ = $7.47, p = 0.006$ which is in line with $H1_a$ and $H1_b$. See Figure 3.4 A) on the next page for a graphical illustration.

How choice architecture affected reading time ($H2_a$ & $H2_b$)

A Kruskal-Wallis test was used to adjudicate whether there was a statistically significant difference between the three groups in terms of the time they spent on the privacy diolog boxes. The three groups had significant differences in the time they spent on the privacy dialog boxes χ^2 $(df = 2, n = 183) = 13.02, p = 0.001$. As a follow-up, Dunn's tests were conducted between two pairs of the experimental groups using a Bonferroni correction. These tests indicated a suggestive difference between PDB1 and PDB2 $(z = -2.84, p = 0.009)$. This result is in line with $H2_a$. However, in relation to $H2_b$ no differences were found between PDB2 and PDB3 $(z = -0.67, p = 1.000)$, see Figure 3.4 B).

How choice architecture affected recall ($H3_a$ & $H3_b$)

Another Kruskal-Wallis test was used to gauge differences between the three groups in terms of their immediate recall scores. When all participants per group are taken together (viewers and non-viewers), the test showed no significant difference in recall scores between the three groups χ^2 $(df = 2, n = 183) = 0.26, p = .879$.

In other words, if all the bars in Figure 3.4 C) are aggregated within each experimental group, no differences are detected. $H3_a$ and $H3_b$ are therefore not supported.

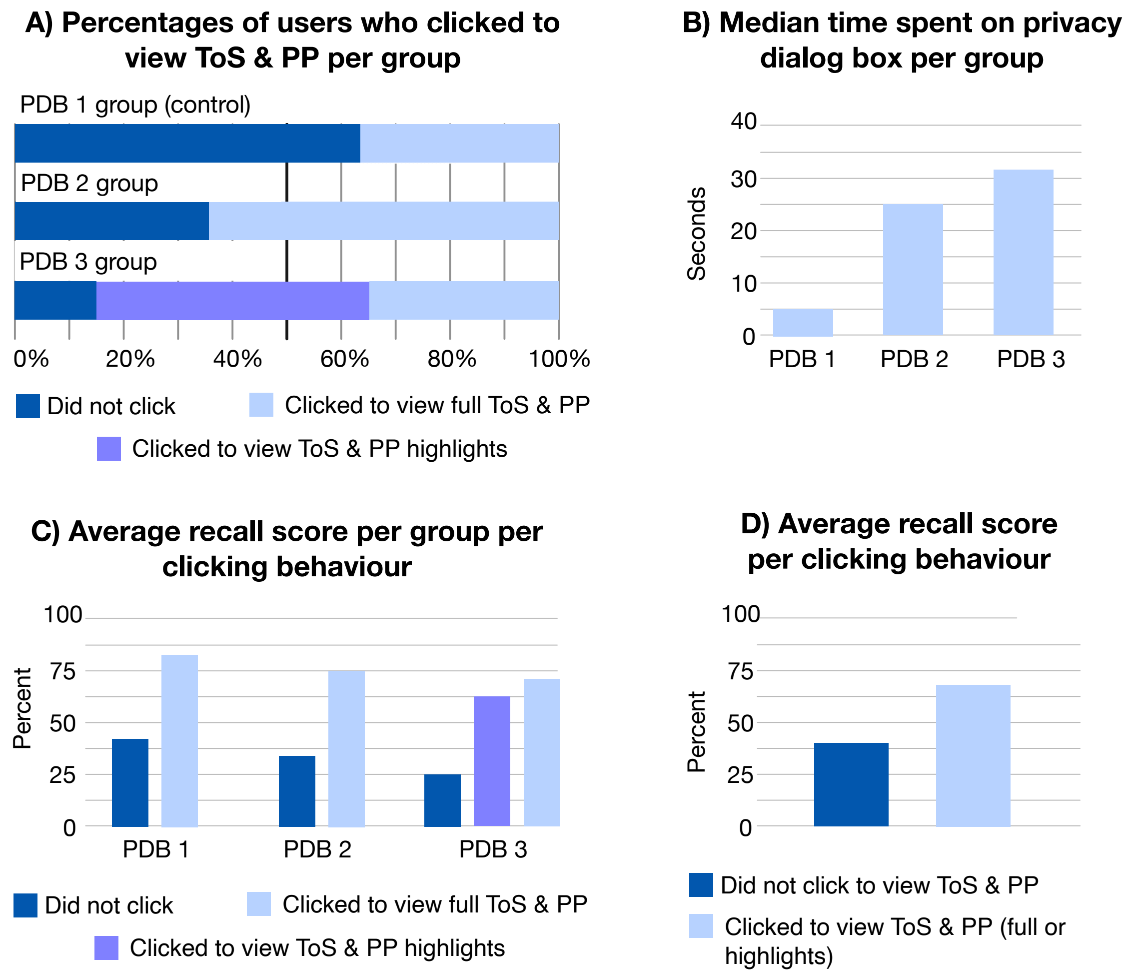

Figure 3.4: Visualizations of the data from the randomized online experiment ($n = 183$)

Yet, a Mann-Whitney U-test revealed a significant difference in recall scores between those that viewed any portion of the terms and privacy policy i.e., Everything or Highlights ($Md = 7$, $n = 118$) compared to those that did not ($Md = 3$, $n = 65$), $U = 1746.5$, $z = 6.14$, $p = 0.000$, $r = 0.45$. Figure 3 D) demonstrates this difference.

3.6 Descriptive Field Study

The IKEA Place app is an example of an iOS/Android mobile application using the choice architecture featured in PDB3. Place is an augmented reality app that lets people virtually place true-to-scale 3D models of IKEA products (e.g. a chair) in their own space (e.g. their kitchen). Even though this augmented reality app does not ask for user identity, it requires access to the camera of the user's device to work properly. Furthermore, as the app is aimed to be used in the user's home, it can potentially access sensitive personal information. Thus, raising the user's awareness about the app's ToS and PP is important to build trust with users. Figure 3.5 shows the UI of the IKEA Place app. When opening the app, users can choose between 3 options just as in PDB3: 'Nothing really', 'Just the Highlights' or 'Everything'.

If they choose option 1, they are directly brought to the agreement page where they have to confirm that they have read, understood and accepted the ToS and PP. If they choose option 3 'Everything', they are shown the full ToS and PP. They can then press continue and they arrive at the agreement screen. If they choose option 2 'Just the highlights', they are presented with the summary of the Tos and PP. The summary take the form of four successive screens, with an image on each screen along with a short explanation. These screens include the camera, improving the app, keeping data safe and your rights as shown in Figure 3.5.

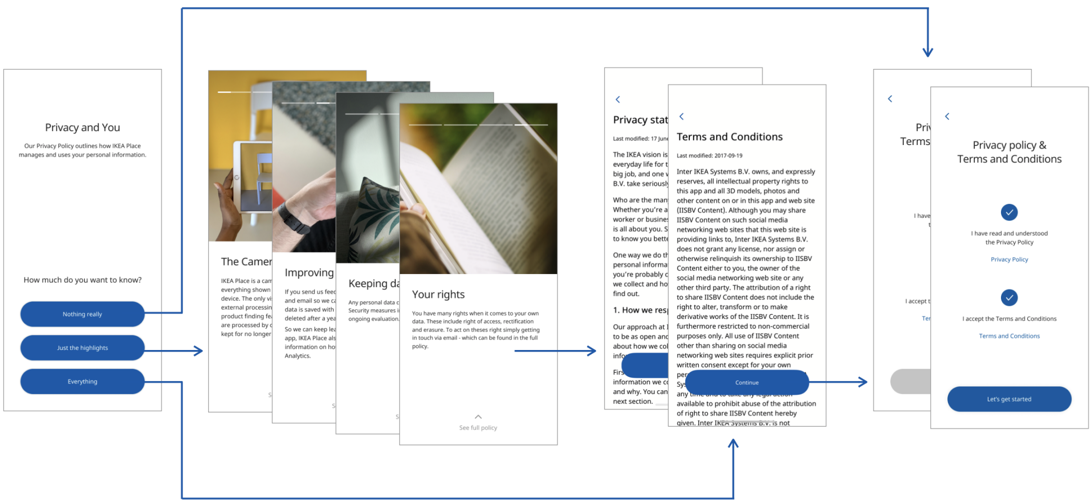

Figure 3.5: The user experience related to online privacy awareness in the IKEA Place app.

3.6.1 Evaluation in the Wild

To gain further insights into how this choice architecture affected actual user behavior, we analyzed anonymous usage data from the IKEA Place app. To do so, we deployed an analytics engine to record activity traces from November 7th, 2019 to November 27th, 2019. During that period, 81'431 users were active on the app. We mainly focused on the option selected by users when they opened the app for the first time. Dropout rates were also recorded during the process, as well as viewing behavior of the full ToS and PP.

How many users viewed the ToS and PP

The results revealed that (46%) of the users chose to read the ToS and PP. Here, Option 1 (Nothing really) is chosen 54% of the times, while Option 2 (Just the highlights) is chosen 20% of the time, and Option 3 (Everything) 26% of the time, see Figure 3.6.

How many users proceeded to accept the ToS and PP

The results showed 97.5% of users who selected Option 1 (Nothing really), proceeded to accept the ToS and PP and started using the service. This percentage was 92.2% for users who selected Option 3 (Everything) and 76.6% for users who selected Option 2 (Just the highlights). The dropout rate includes users who abandoned the app and those who went back to the privacy dialog and potentially selected another option. These descriptive results might indicate that the implementation of Option 2, which forces users to stay a fixed amount of time on each one of the highlight screens, might add too much friction to the user experience. Further analysis of the user behavior of those who selected Option 3 indicates that only 11% of users scrolled through 50% or more of the ToS and PP.

These results further illustrate that bringing users to view the ToS and PP is a necessary but not sufficient action to raise privacy awareness.

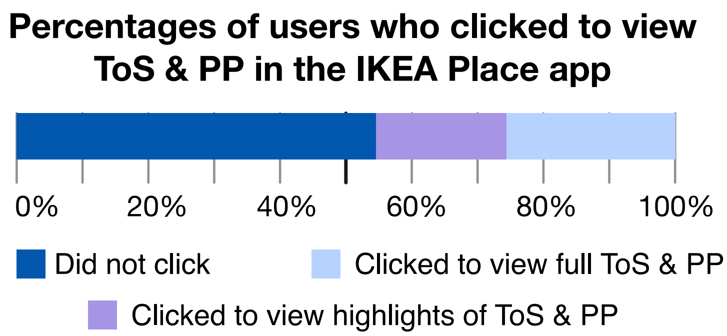

Figure 3.6: Data visualization related to the IKEA Place app users ($n = 81431$)

3.7 Discussion

In this chapter we investigated how the design of new choice architectures could mitigate what is called the biggest lie on the Internet: The fact that people confirm they have read the ToS and PP without doing so. As [235, p. 153] argued, "privacy notices are incredibly detailed legalese designed to indemnify the firm, not protect the consumer". As a result, users lack motivation and ability to read this content. The problem is aggravated by service providers using the quick-join choice architecture, which encourages users to ignore the ToS and PP by default. To tackle this problem, we used two digital nudges to design two privacy dialog boxes: One aimed at sparking users' motivation by throttling mindless activity with an explicit choice (PDB2) instead of having an opt-in default (PDB1). Similarly, the other privacy dialog box was designed as an explicit choice but also aimed at facilitating the user's ability by providing a meaningful summarized alternative (PDB3). Our experimental results show that with both these designs, the number of users who view the ToS and PP is significantly higher than with the quick-join design. Interestingly, when we compare the quick-join standard with the PDB3, adding the (highlights) option converts non-readers into highlight readers but does not lead those who read the full ToS and PP to only read the highlights when that alternative is suggested, see Figure 3.4 A). Our results converge with prior research showing that framing the choice as a neutral question (PDB1 vs. PDB2) rather than an opt-in policy can drastically change users' choices [224, 225]. Furthermore, the results show that suggesting a simpler alternative also drives viewing behavior in the context of privacy awareness. This design consideration has worked in other contexts as well [226, 227, 228].

However, the results of this study are mixed. Although a strong case can be made that viewing and reading the ToS and PP should increase participants' recollection of that content, neither one of the two choice architectures changed the average recall score. Based on the description of the data in Figure 4 C), we observe that users who read the ToS and PP in PDB1 tended to recall it better than those who read the ToS in PDB2 and 3. The same pattern is observed for those who completely ignored the terms and privacy policy. Thus, although PDB2 / - 3 vs. the quick-join group (PDB1) increase the proportion of users who view/read the ToS (or its highlights), those who read it, recall it less accurately. The additional users who read the ToS and PP due to the new choice architectures may be users who are less interested in the actual content (compared to those who read the ToS and PP from the quick-join environment). Although they click and read because of the nudges (which probably entails a quite automatic processing of the information), they might only read the content superficially, using limited cognitive resources. A possible risk in using nudging techniques that tap onto the automatic mind, is their lack of any educational effects [236]. This could explain the unchanged recall scores in the nudged groups.

Reading is an important outcome because many users confirm they have read and agreed to ToS and PP without actually having read them, which engages their legal responsibilities towards online service providers. In the context of future research it is also worth asking whether recall is an important outcome. If the aim is to move users towards more informed consent of the ToS and PP, then we argue that recall seems to be a good proxy measure. An alternative aim for future research could be to encourage users to act based on this knowledge. An example would be by changing the privacy settings of the actual app either via the application itself or a plugin. Another future research direction could be to investigate whether user actions are contingent on their reading or recollection of the privacy policy (or on none of them).

The implications of this study are salient at a time when technology allows for new types of data to be collected without much user awareness [190, 237]. These results provide an initial step to raise users' awareness by nudging them to view and read the relevant information related to their data privacy (even if they do not necessarily recall it better). If regulation and policy does not sufficiently assist or protect consumers, will service providers have an interest to increase privacy awareness among their users? One answer could come from organizations which strive to establish trust with their customers. The IKEA Place app is one example of a service provider willing to raise users' privacy awareness. If users become more aware of how their data is treated by each service provider, they can demand that these services become more sensitive to online privacy. The type of data that is collected and the way it is used could become part of the value proposition together with the service itself.

3.7.1 Limitations

Our proposed implementation of the digital nudges represents one instance among many possible alternative designs, which implies that our solution can be further optimized by both researchers and practitioners. The following limitations should be noted. First, our data from the IKEA Place app is only used for a descriptive purpose to show that that these digital nudges can be readily implemented by large online service providers to facilitate the informed consent of their users. We cannot validate that this design is superior in this larger context due to the lack of comparable control groups, we only do this comparison in the randomized online experiment. Second, the MTurk population was compensated to conduct the task we proposed. Some participants might have felt that it was part of their task to read the ToS and PP even if we did not make any such suggestions.

Indeed, compared to real users of the IKEA Place app, a higher proportion of online experiment participants viewed both the highlights and the full ToS and PP. Prior scholars have suggested that the Mturk population can be more attentive than average students [238]. However, differences such as these do not change across the groups of the online experiment. Third, we cannot control for novelty effects in our findings i.e., that novel choice architectures such as in PDB2 - / 3 facilitated viewing and reading behavior by the fact that the participants had not seen this particular design before. Yet, we did take steps during the analyses of the online experiment by following the recommendations of Benjamin et al. [234] to reduce the rate of false positives for new findings.

3.8 Conclusion

With more and more personal data being gathered and processed online, raising the bar for 'informed consent' is an important design challenge. Today, companies are legally compelled to inform their users. Yet, few users are informed when they consent to how, why and where their personal information is processed. In this chapter we have shown that HCI and system designers can affect the privacy awareness of users through digital nudges that increase the likelihood that the ToS and PP are read.

The proposed design that implemented two nudges (PDB3) showed around a 75% decrease in number of users who just agreed the ToS and PP without even viewing them. This significant reduction in users who committed the so-called biggest lie on the Internet is a promising start. While the results showed that clicking the ToS and PP did increase the users recall of the contents, the proposed designs did not inherently increase recall among users. The field study with over 80'000 users of the IKEA Place app demonstrates that these digital nudges can be readily implemented by large companies. Yet, there is still a lot of work and research to be conducted before we can generally claim that users have given their informed consent online.

4 Reducing a Newsfeed for Social Media Detox

A version of this chapter has been published as: Aditya Purohit*, **Kristoffer Bergram***, Louis Barclay, Valéry Bezençon & Adrian Holzer. *Starving the Newsfeed for Social Media Detox: Effects of Strict and Self-regulated Facebook Newsfeed Diets.* 2023. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). ACM, New York, NY, USA.

*Both authors contributed equally to this research

Contents

4.1 Introduction

In this chapter we will explore a more interconnected and personalized digital choice architecture, namely, a social media newsfeed.

In some countries, the self-reported averages of social media use are 3–4 hours a day [239]. The large-scale societal consequences of increased social media consumption are extensive. They range from more positive use-cases such as habits [240, 241], to the negative such as mental health problems [242, 243] and degraded social interaction [244]. At the heart of this issue lies compulsive social media use which is a "repetitive, ritualistic behavior involving an individual's inability to control, reduce, or stop the use of mobile [social media]" [245, p. 177].

Each user is potentially vulnerable to harmful use-cases through various reinforcement cycles when using social media platforms [245, 246]. During such reinforcement cycles, interactive technologies serve as a form of stimulus, where compulsive usage becomes a response and the reinforcement cycle continues. Social media platforms are designed to maximize connections and time spent online leading to hyperconnectivity by employing user interface designs, sometimes referred to as dark patterns [247, 248, 249], i.e., designs that utilize knowledge of human psychology to trick individuals into an act that is not in their best interest [51, 249].

HCI researchers have previously pointed to Facebook's *infinite newsfeed* as a dark pattern [249]. The infinite newsfeed taps into three main characteristics of dark pattern as defined by Roffarello et al. [249]: (1) the design undermines individuals' autonomy and distracts them from their goal [249], (2) individuals experience that time has passed and they have no control over it [249] and (3) in retrospect, the design makes a person regret the time spent on the service [51]. In addition, the infinite newsfeed negatively influences individuals' digital wellbeing [250]. As such, the newsfeed is likely to be among the factors that eventually lead people to use Facebook compulsively. In this chapter we aim to better understand how newsfeed is related to time spent on Facebook and how a digital newsfeed diet, i.e., an intervention that reduces one's newsfeed, could be designed to reduce Facebook usage and potentially improve digital wellbeing. This leads to our overarching research question.

RQ: How does a newsfeed restriction affect Facebook usage?

This overarching question leads to three sub-questions. First, how do different types of newsfeed restriction (abstinence vs moderation) affect time spent on Facebook? Second, how does newsfeed restriction affect user experience? Third, how does newsfeed restriction affect time spent according to users' compulsion to use Facebook? In this chapter, we attempt to answer these questions by taking a design science research methodology (DSRM) approach [251, 252, 253, 254, 145]. In DSRM, the artifacts that impact people and organizations are created and evaluated in order to solve identified research problems [212] in six steps. This chapter is structured following these steps. The introduction covered the first step, namely the definition of the problem. The second step, i.e., defining the objective of the problem, is covered in Section 4.2, which lays out our hypotheses. The third and fourth steps, i.e., design of the solution and demonstration, are covered in Section 4.3. The fifth step, i.e., the evaluation of the solution, is split between Section 4.4, which presents the evaluation setup, and tests the hypotheses, and Section 4.5, which provides qualitative insights into the user experience through an exploratory thematic analysis. Finally, the sixth step, i.e., communication of the findings, is covered in the discussion in Section 4.6 and the conclusion in Section 4.7.

4.2 Related Work and Research Objectives

Here, we describe relevant prior work, which led to the six specific research hypotheses that we tackle in this chapter. In particular, we review work that relates social media use to compulsive use, before we give an overview of attempts to redesign the newsfeed to address this issue and how this could relate back to compulsive use and user experience.

4.2.1 Social media and compulsive use

There has been a lively debate among scholars about where chronic social media usage ends and usage more akin to "addiction" begins [255, 256, 245]. However, there seems to be substantial agreement among HCI researchers that social media platforms are designed in a way that facilitates compulsive use [257, 256, 258, 179, 259].

Typically, compulsive social media use refers to when the user is unable to self-regulate how frequently they use a given platform or how much time they spend on it [260]. Compulsive social media use is an important societal concern because users facing it might have psycho-social and professional consequences [261, 245]. We previously mentioned that other researchers have theorized that compulsive social media usage happens through a series of reinforcement cycles, with interactive technologies serving as a form of stimulus, and compulsive usage becoming a response so that the reinforcement cycle continues. We posit that the newsfeed on Facebook is one such interactive technology. While several previous studies have confirmed a strong correlation between a variety of compulsive use measures and self-rated time on social media [262] – we want to validate this assumption by examining the relationship between compulsive use measures and actual time spent (by real users) on a well-established social media platform: Facebook. Previous scholars have also argued for the importance of validating the association between compulsive use and actual user behavior on social media platforms [245, 263]. Therefore, we hypothesize the following:

H1: Compulsive use is positively correlated with actual time spent on Facebook.

4.2.2 Redesigning the newsfeed

An important design feature for prolonging usage time on social media platforms comes through the newsfeed [264]. While the feature itself is the principal entry-point for a Facebook user, a recent survey suggested that it was also one of the features that tends to be associated with regret for the user [51]. Several previous studies have suggested the removal or reconfiguration of the Facebook newsfeed as a potential context for design researchers to tackle. Suggestions in the HCI literature range from complete removal [265], contextual removal that is dependent on the given user's goals [266], to adding various filtering options [266, 259, 179]. For example, both scholars and sampled users have made suggestions around configuring the newsfeed to only show content from close personal friends [259, 179]. In another study, Lyngs et al. conducted a user experiment where different Chrome extensions were used to compare a control group to goal reminders and a completely hidden/blocked newsfeed [266]. Here, a majority of sampled users suggested that they wanted more granular control over the Facebook newsfeed. In short, this newsfeed blocking intervention was effective when it came to decreasing users' visit length on Facebook's site but also led to fear of missing out (FOMO). Another study concerned the design and evaluation of a Chrome extension that allowed to hide some parts of the Facebook user interface or remove colors to make it less appealing [179]. When considering how users spend their time and attention across apps, previous researchers have characterized design interventions along a spectrum ranging from internal to external [259]. External mechanisms entail monitoring problematic apps from the outside, and informing or notifying the user, e.g., setting a maximum duration on an app, or viewing the time spent on an app. Other external examples include the use of the phone's vibration to signal to the user that they have spent enough time on Facebook [102], framing the feedback negatively or positively [267], locking users out of apps [258] and locking users out of the device [268]. Internal design mechanisms would entail redesigning certain features of a problematic app directly inside the app, e.g., removing the newsfeed. Such internal mechanisms might remove more problematic aspects from a given app, while still retaining its most important benefits [259]. HCI researchers have highlighted that Google Play, Chrome Web and Apple App stores have created a vast marketplace for tools that help users in their online struggles for self-control (see [269, 270] for reviews of such tools).

In this research context we aim to change an internal design mechanism, more precisely the dark pattern such as infinite newsfeed that leads to mindless scrolling [249, 271]. We argue that by designing an artifact that restricts the newsfeed, that is, that reduces the content of the newsfeed from friends, pages and groups, users would be less exposed to variable rewards [179, 272] thereby breaking the inexhaustible reinforcement cycles of anticipation, uncertainty and feedback. We provide two versions of the newsfeed restriction artifact: (1) a *strict newsfeed diet* that reduces the newsfeed to the minimum and (2) *a self-regulated newsfeed diet* that reduces the newsfeed to a minimum, but lets users fill it with updates from people, pages and groups that the users opt-in to follow (as opposed to the system's default setting to follow all connections). In both cases, we hypothesize that the restriction will reduce time spent on Facebook:

H2a: A strict newsfeed diet will decrease time spent on Facebook.

H2b: A self-regulated newsfeed diet will decrease time spent on Facebook.

4.2.3 Newsfeed diets and compulsive use

Previous research investigating various digital self-control artifacts hinted that such tools seemed especially useful for users who experienced the most intense struggles for self-control or labeled themselves as addicted [269]. This finding from sampling public user reviews can be contrasted with previous research on habits. Here, literature points to the general difficulty of changing strong (vs weak) habits [273]. That is, stated intentions are at odds with actual behavior in the presence of strong habits, and this is especially salient when the environment is stable, because the cues triggering the habitual response typically come from the environment [274]. In the case of newsfeed restriction diets, the environment stays stable; only the content of the newsfeed changes. Therefore, we expect more compulsive users to trigger their habitual response (i.e., spend time on Facebook) whatever the content and the restrictions applied to their newsfeed. This should be less the case for less-compulsive users, whose time on Facebook should be more affected by newsfeed restrictions. This is reflected in other, digital behavioral interventions, such as digital nudges, which tend to be less effective when habits are established [74].

H3: The newsfeed diets will affect less- (vs more-) compulsive users' time spent on platform to a greater (vs lesser) extent.

4.2.4 Newsfeed diets and user experience

Finally, we believe that a potential consequence of imposing a *strict* newsfeed diet, where users have no say in how the newsfeed is restricted, may adversely impact the user experience. While this intervention is not expected to affect usability components such as ease of use, users might potentially feel FOMO [275, 266], or difficulty to connect with friends and family. One way to address potential issues related to FOMO or connecting with important people in one's life is to take a self-regulation approach to the newsfeed instead of a strict restriction. A recent study that focused on increasing users' sense of agency on Twitter with internal vs external design mechanisms found that only internal mechanisms significantly increased users' sense of agency [276]. Findings from the same study suggest that usability issues were most frequently mentioned when users feel as if they are not in control [276]. The self-regulated newsfeed diet lets users define which connections to follow. In doing so, it gives users a sense of control, which is important to user satisfaction [259, 277, 276]. Moreover, HCI has long emphasized the importance of a sense of control over how users experience the interaction with technology [278]. The users want to feel that they are in charge and that the system responds to them [278]. This leads us to the following hypotheses:

H4a: A strict newsfeed diet (vs no diet) will negatively impact user experience.

H4b: A self-regulated (vs strict) newsfeed diet will positively impact user experience.

4.3 Design and Demonstration

The problem statement pointed out that social media users tend to mindlessly spend too much time on social media because of the infinite newsfeed feature. The solution's main objective is to address this issue by reducing the newsfeed.

4.3.1 Newsfeed mechanics and design

Social media platforms use different techniques to populate the user's newsfeed. Even though the exact algorithms are not open to the public, they follow certain principles. At the center of such algorithms is the notion of following some content emanating from friends, pages or groups. When a user follows a certain contact – the user's newsfeed will contain updates from that contact. It should be noted that different social media platforms have different ways to call similar types of relations. Generally, social media platforms offer at least two types of relations: a relation that allows users to establish a contact with someone (called 'friends' in Facebook, 'contacts' in LinkedIn, or 'followers' in Instagram) and a second one that allows to see or not see someone's update in one's newsfeed (called 'follow' or 'unfollow' in Facebook and LinkedIn and 'unhide' or 'hide' in Instagram). By default, these platforms generally establish the second relationship, i.e., showing people's updates in one's newsfeed, when the first relationship is established. In this chapter, we use the term *contact* for the first relationship and we use the terms *follow and unfollow* for the second relationship. As such, the newsfeed is a rough function of the number of contacts that a user follows. In addition to updates from followers, a Facebook user's newsfeed is also populated with different sponsored posts, as most mainstream social media platforms are monetized through advertising.[*] In the context of Facebook, sponsored posts are personalized for users, based on different information such as stated preferences, demographics, location or previous online activity [280].

In some solutions, a browser extension can directly hide parts of the UI such as the newsfeed [179, 266]. They can do this by simply modifying the HTML/CSS code on the client side. However, this type of solution has a limitation that makes it

[*] In 2021, Meta generated more than 99% of their total revenues (USD 115.655 billion) from advertising [279]

unsuitable for this study: The limitation is the fact that modifications are restricted to browser-based solutions and not applicable to native mobile apps, which is how many social media platforms are also used today. With the number of devices that allow access to social media content, data from a single device alone may not suffice to capture people's digital behavior [281]. One way to offer a cross-device solution is to provide an internal intervention that will, by definition, spread to all devices. One such approach to limit the newsfeed is simply unfollowing certain contacts. To unfollow means to stop seeing these contacts' posts in one's newsfeed. Unfollowing a contact does not remove you from each other's contacts list and it still allows you to go see your contacts' updates on their profiles. If a user unfollows every single contact, their newsfeed will be mostly empty, except for potential sponsored content.

To unfollow everything, users can navigate to a specific contact, but in some platforms it can be hard to find this unfollow option. To unfollow a contact, a Facebook user must currently navigate to the contact's profile, hover over *Following* (on a friend's profile) or click the *Following* button (on a page or in a group) near their cover photo and then select the *Unfollow* option [282]. This action then has to be repeated for each contact. Whereas the action of following a contact is the default state of affairs when a user adds a new contact, the unfollow procedure seems to suffer from lack of discoverability, which could be considered an example of another dark pattern [37]. With users following an average of over 300 friends, pages and groups on Facebook, repeating this manual procedure can also quickly become tedious [283].

4.3.2 Designing an automated unfollow solution

Automating the unfollowing process for large numbers of contacts could ease the process and could also address the cross-device issue. To enable this process to happen in a user-friendly manner, we designed a Chrome web browser extension that allows the user to unfollow all friends, pages and groups that they follow on one particular mainstream social media platform, namely Facebook. This extension was able to be downloaded and installed directly from the Chrome Web Store. From a user's perspective, the experience occurs as follows: once installed, the extension automatically unfollows a user's friends, pages and groups from Facebook, when the user opens Facebook via the Chrome browser, without any user interaction needed. Technically, the extension accesses the list of friends, pages and groups of a user through an authentication token. It then iterates through the list and performs POST requests for each friend, page and group to unfollow them. It should be noted that the process is not immediate, taking about 5–15 minutes to unfollow 300 friends, pages and groups (longer for users with many more friends, pages and groups).

By default, users are in what we are calling the '*self-regulated*' condition. That is, the extension unfollows all of the user's friends, pages and groups, and then the user can refollow anyone they please. For the purpose of the field experiment, we modified the extension in two ways: First, we added an extra prompt after the extension was installed and before it started unfollowing friends, pages and groups, in order to separate regular users and study participants, see Figure 4.1. The existing extension prompted regular users to directly go to Facebook after the installation, whereas study participants were instructed via an extra pop-up window to click a link so that they could enter a randomly assigned unique ID that they were given via a Qualtrics survey; this enabled their online behavior to be linked to their survey responses. Second, we modified the extension to assign study participants to three conditions based on their unique ID: 1) The default self-regulated diet, 2) A control condition in which the extension did not trigger any unfollow process, 3) A strict diet. To put users on the strict diet, we programmed the extension in such a way that it would unfollow all friends, pages and groups. If a new friend, page or group was added, or if an old friend, page or group was refollowed, the extension automatically unfollowed them. Finally, in order to allow users who use the extension to get their newsfeed back to normal, we designed a function to refollow everyone. Similarly to the unfollow process, this function iterates through all of the user's friends, pages and groups and refollows them one by one.

In addition to the core functionalities of the solution, we also designed infrastructure to measure the time spent on the Facebook platform for the purpose of the study on the site itself as well as on the iOS app. To measure time spent on the site (we focused on the Chrome web browser), we designed the web extension so that it would send time spent for each session to an analytics server. The extension typically performs a check every second to see if the browser's active tab is on the social media platform. If the extension detects a visit, a timestamp is created with an initial duration of 0 seconds. Then for every subsequent second on the site, the duration is incremented and sent to the analytics server.

To measure the time spent on the app (we focused on iOS devices), we implemented an automation process on the Apple shortcut app similarly to the method used by Purohit et al., [272, 284]. In a nutshell, the Apple Shortcut app allows the design of an automation that is triggered when a particular third-party app on the phone is opened or closed. The automation then writes timestamps, every time the Facebook iOS app was opened and closed, in a CSV file that was stored locally on each user's phone. At the end of the study we simply asked users to submit this CSV file to inspect the time they spent on the app without the need to build a client server infrastructure for the mobile data.

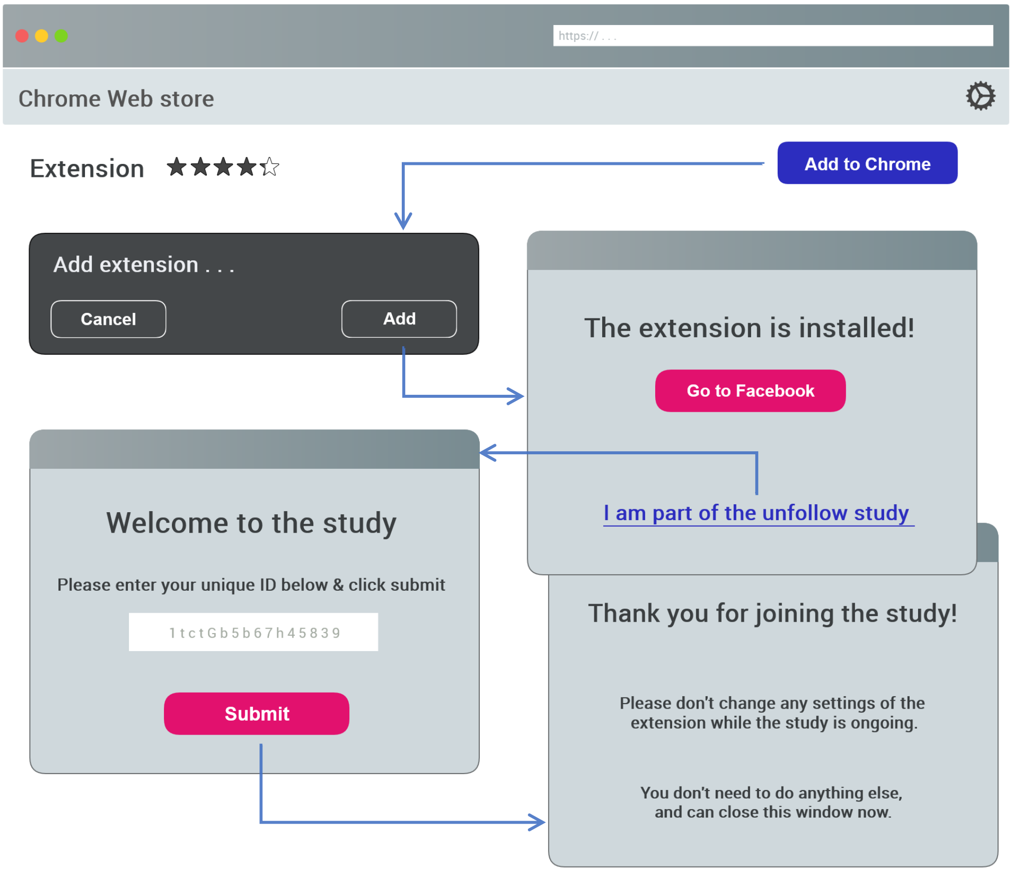

Figure 4.1: Mockup showing the installation of the extension: After adding the extension to the browser, three pop-up windows appeared where users (regardless of which condition they belonged to) had to click a link, input a unique code, and then close the final window.

4.4 Evaluation

The evaluation is divided in four main phases: (1) screening survey, (2) main survey, (3) the field experiment and (4) the exit survey.

First, participants filled out a short screening survey to confirm that they were interested in and eligible for the study. Owing to technical design choices, only Facebook users who accessed the platform with the Google Chrome browser and the iOS app were eligible. At that stage, it was already made explicit that their continued participation in the study could entail unfollowing all of their contacts on the platform.

Second, after this screening, a main survey was conducted. Users were asked to share demographic information, and fill out usability and compulsive use scales. Then they were asked to download, install, and configure the Google Chrome extension and Apple's Shortcuts app. As a verification procedure, each user received a random six-digit code to establish a link between the survey data and the behavioral data from the browser and the mobile app. During this installation, users were randomly assigned to one of the experimental groups using Qualtrics.

Third, the field experiment was performed. As mentioned previously, the field experiment had three randomized groups: no newsfeed diet (control), the strict newsfeed diet condition, and the self-regulated newsfeed diet condition. The field experiment was divided up as a seven-day baseline period, followed by a 14-day treatment period. During the baseline period, the extension remained idle for all three groups. The idea was to measure baseline daily average time on Facebook's site/app for each user. At the beginning of the treatment period, the different conditions were activated. For users in the control group, the extension was not activated and these users were told that they could continue to follow all their contacts. For users in the strict newsfeed diet condition, the strict version of extension was automatically turned on and it was communicated to this group of users that they would not be able to follow their contacts during the following 14 days. Similarly, for users in the self-regulated newsfeed diet condition, the extension was automatically turned on, but in this case in its self-regulation version which allowed users to refollow any contact they wanted. Users were again told explicitly about what was going to happen, i.e., that all their contacts would be unfollowed, but that they were able to refollow them during the succeeding 14 days if they wished to do so. As the Chrome extension's unfollowing procedure took around one hour for a thousand contacts, it could potentially take a long time for users with many contacts. We considered that the process of unfollowing contacts might take several days if users did not spend enough time on Facebook to unfollow all contacts on the day they installed the extension. To minimize interference of this process with their regular usage, we only use the behavioral data of the latter seven days of the 14-day intervention as treatment period data.

Fourth, at the end of the third week, participants completed an exit survey that contained the open user experience questions and they were shown how to upload their mobile usage data (the CSV file described above) and finally they were shown how to delete the Shortcuts automation and refollow their contacts if they wished to do so.

4.4.1 Study design and data overview

The Sankey diagram in Figure 4.2 summarizes the overall design of the study and the most important sources of data. The size of each pipe and node is proportional to the number valid responses, i.e., the size of the sample. The Sankey diagram is ordered from left to right in terms of sample attrition rate. That is, the pool of valid cases decreases from left to right in Figure 4.2. The data sources in Figure 4.2 are not necessarily in chronological order. In terms of chronology, the behavioral data

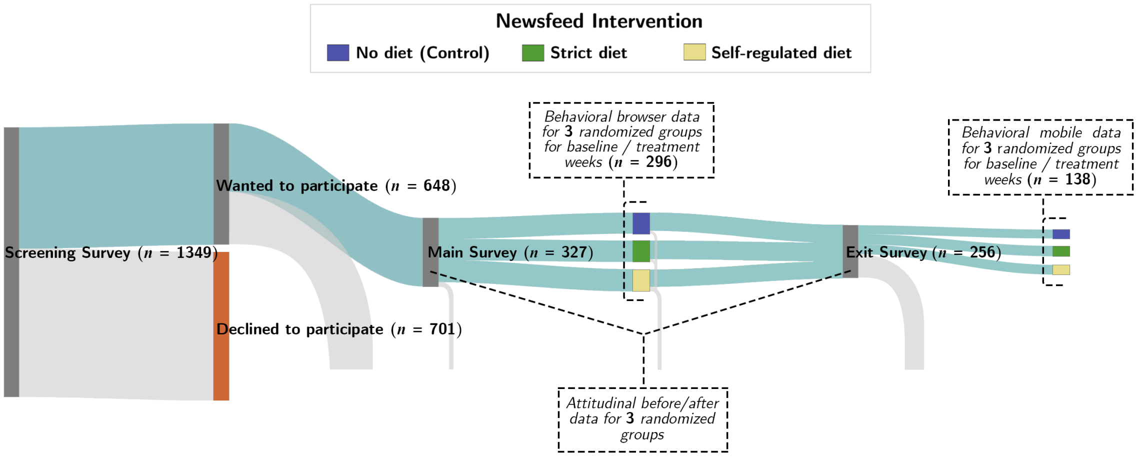

Figure 4.2: Sankey diagram [285] illustrating relevant data sources for the study, together with the sample attrition rate.

collection (both mobile and browser) started right after the main survey and up until the exit survey. During the exit survey, the remaining users were asked to upload their CSV files from their mobiles, which is why the sample attrition rate increases after the exit survey.

The initial sample size for the screening survey was n = 1349 US Facebook users. In total, 327 users enrolled on the study by completing the main survey. Out of those, 296 users successfully downloaded/installed the Chrome browser extension. Just 256 users finished the exit survey and thereby provided valid answers to the attitudinal before/after variables. From these, we received valid mobile data from 138 users in total. The last number constitutes the final number of users for the study and, as the next section will show, this number slightly fluctuates, depending on the variable of interest. Despite the attrition rate in the study, the three randomized groups remained fairly balanced throughout the field experiment.

4.4.2 Results

Table 4.1 presents the descriptive statistics of variables of interest collected at the *baseline period* of the field experiment. This is mainly to ensure that the randomization procedure worked during the field experiment so that the intervention groups are comparable. Table 4.1 shows the sample medians ($\hat{m}$), averages ($\hat{\mu}$), standard deviations ($\hat{\sigma}$) and the number of valid cases (n) for each variable across the intervention groups. The high standard deviation for time on site/app shown in Table 4.1 reveal a very high variation of Facebook usage between users. Indeed, the measure ranges from roughly seven seconds to almost four and a half hours per day. Nevertheless, the average of 35.7 minutes per day that we observe is in line with figures of the Global Web Index report 2021, which shows that the average monthly time spent on Facebook is 19.5 hours, i.e., 39 minutes per day [286].

Table 4.1 also highlights that there seems to be a higher sample attrition rate in the control group compared to the treatment groups. While the control group is smaller, a χ^2goodness-of-fit test indicated that there were no significant differences in the proportion of valid cases in the three randomized intervention groups (42, 49, 47) when compared to the expected proportions of (46, 46, 46), $\chi^2(df = 2, \ n = 138)$ =

Table 4.1: Descriptive statistics across each newsfeed diet intervention group during the baseline period of the field experiment.

	Control				Strict diet				Self-regulated diet			
Variable	$\hat{m}$	$\hat{\mu}$	$\hat{\sigma}$	n	$\hat{m}$	$\hat{\mu}$	$\hat{\sigma}$	n	$\hat{m}$	$\hat{\mu}$	$\hat{\sigma}$	n
Time on site/app (s)	1385.9	2044.6	2650.1	42	1594.7	2087.5	2020.2	49	1090.3	2279.9	3069.9	47
Compulsive Use	3.750	3.953	1.384	43	3.625	3.730	1.350	50	3.250	3.702	1.910	47
SUS score	80.00	80.23	14.19	43	75.00	73.50	18.70	50	80.00	78.03	13.61	47
Age	29.0	31.05	10.51	43	28.0	31.28	11.95	50	24.0	26.89	7.98	47
Gender (% female)		65.12		43		58.00		50		63.30		47

0.565, p $= 0.754$.

In fact, Kruskal–Wallis tests and another χ^2-test detected no significant differences between the intervention groups on any of the listed variables in Table 4.1. We therefore conclude that the medians and proportions of these variables are reasonably balanced across the intervention groups at the baseline period of the field experiment.

Time on platform, compulsive use, and newsfeed diets (H1–H3)

When performing tests for statistical significance, we followed the guidelines outlined by Benjamin et al. [234]. The relationship between time on site/app during the baseline period and compulsive use showed a significant correlation: $\rho = 0.390$, [95% CI: $0.233, 0.526]$, $n = 138$, $p < 0.001$. That is, high levels of time on site are associated with higher degrees of compulsive use. H1 is supported.

To conform to the assumptions of parametric models and tests, we applied a natural log transformation to the daily time on site/app variable. The results before and after this log transformation are visualized in Figure 4.3. The raincloud plots in Figure 4.3 combine the distribution curves, boxplots for indicating where the density of the distribution lies and the scattered rain of all the individual data points [287]. The colored lines highlight the mean differences between the intervention groups. Figure 4.3 is complemented by Table 4.2 which also shows the sample size, mean and standard deviation from each intervention group after the log transformation. Table 4.2 also highlights that we were not in complete control to enforce the newsfeed diets. We captured partial data on the percentage of unfollowed contacts which suggests that the unfollow procedure was not completed by all users in the field experiment. This refers to average and median percentage changes in followed contacts between baseline and treatment periods for each newsfeed diet. Table 4.2 highlights large differences between the control and the newsfeed diets, but not between the two newsfeed diets. We

Table 4.2: Log transformed time on site/app measures for the baseline and treatment periods across each newsfeed diet intervention. Missing values in the percentage change of followed contacts across each newsfeed diet: No diet: 16, Strict diet: 19; Self-regulated diet: 24.

	Baseline period			**Treatment period**			**% Change in followed contacts**	
Newsfeed intervention	n	$\hat{\mu}$	$\hat{\sigma}$	n	$\hat{\mu}$	$\hat{\sigma}$	average	median
No diet	41	7.040	1.220	41	7.039	1.515	↓ 3.3	↓ 0.0
Strict diet	47	7.175	1.174	47	6.157	1.544	↓ 62.0	↓ 82.7
Self-regulated diet	46	7.006	1.309	46	6.489	1.560	↓ 65.1	↓ 77.4

also computed the log changes in time on site/app between the baseline and treatment periods.

These log changes are again visualized with raincloud plots in Figure 4.3 D). The untransformed changes in daily time on site/app across each newsfeed diet group can be found in Figure 4.3 C). The connection between the raincloud plots in Figure 4.3 is the following: The distribution curves, boxplots and scattered rain shown in C) are simply the result of subtracting the baseline week data from the treatment week data in plot A). With the log transformed data in Figure 4.3 B), the same subtraction is performed which results in plot D)[†]. To examine the change that the newsfeed diets may have had on time on site/app and to investigate whether restricting the newsfeed had a different impact on users who need it most, i.e., compulsive users, we specified three multiple regression models in a hierarchical fashion. In each model, we regressed the log changes in time on site/app between the baseline and treatment periods on a number of predictors. The dependent variable used for these models is visualized in raincloud plot D) in Figure 4.3. This dependent variable is the natural log change in time spent on site/app between the treatment and baseline periods. In the first model, we simply introduce the two newsfeed diets as dummy variables with

[†]The x-axis of this raincloud plot excludes one outlier at +26145 seconds that can been seen at +2.539 in raincloud plot D).

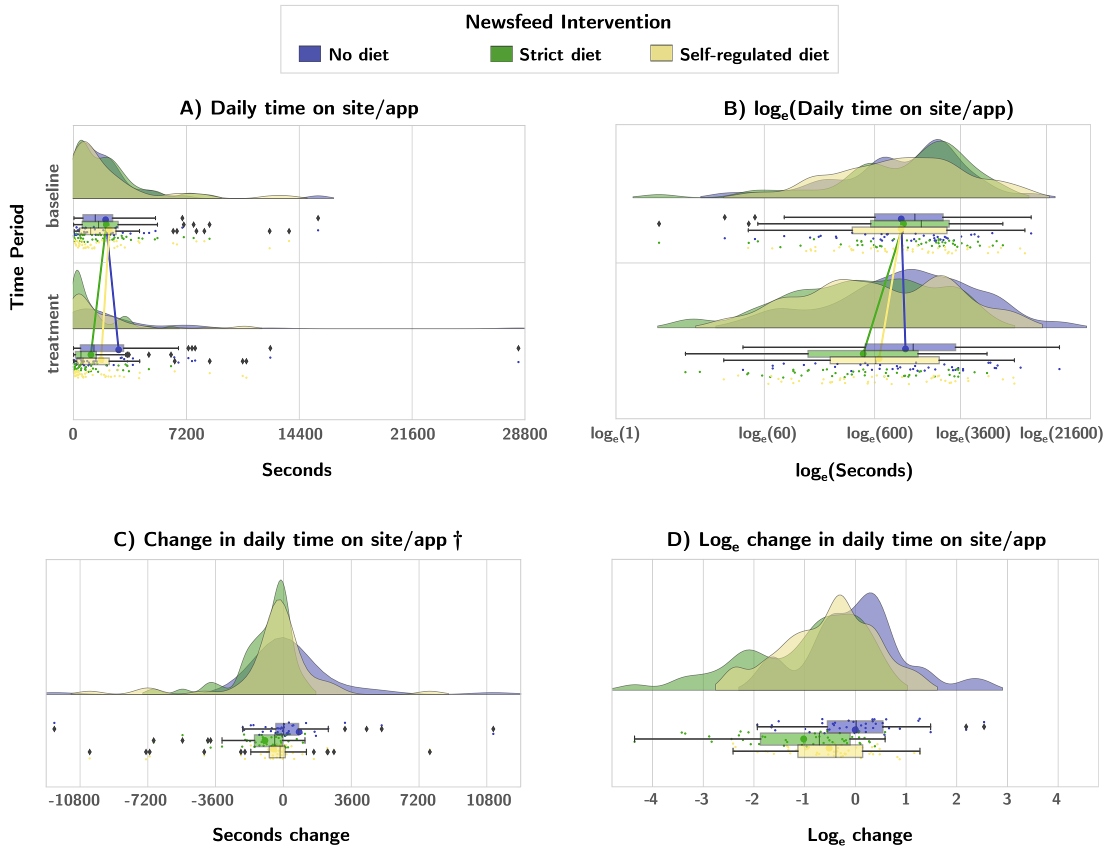

Figure 4.3: Raincloud plots [288] A) and B) showing time on site/app across each newsfeed diet during baseline and treatment weeks, together with C) and D) showing changes in time on site/app between baseline/treatment weeks across each newsfeed diet.

indicator coding. This model is related to H2a and H2b.

In the second model we do the same while also adding users' logged baseline measures for time on the site/app together with their compulsive use measure before the baseline period. The second model therefore accounts for the potential effect from the interventions and compulsive use, and with the fact that different users spent different amounts of time on Facebook during the baseline period. This is sometimes referred to as a conditional change model [289].

In the third model we keep the previous predictors while also introducing the moderating effects of compulsive use during the baseline period on each intervention group e.g. d_1x_2 and d_2x_2. This model relates to H3. The Appendix contains a description of all variables used in the three regression models.

Table 4.3: Multiple linear regression models: Predictors for log changes in time on site/app between baseline and treatment periods.

		Regression coefficients					
Model	Variable	$\widehat{\beta}$	95% CI	Std. Error	t	VIF-value	p-value
1	Constant (β_0)	0.001	[-0.320, 0.319]	0.161	-0.005		.996
	Strict diet	-1.018	[-1.455, -0.581]	0.221	-4.606	1.394	.000***
	Self-regulated diet	-0.516	[-0.955, 0.077]	0.222	-2.323	1.394	.022*
2	Constant (β_0)	0.008	[-0.326, 0.310]	0.161	-0.050		.960
	Strict diet	-0.982	[-1.417, -0.546]	0.220	-4.460	1.403	.000***
	Self-regulated diet	-0.498	[-0.935, -0.061]	0.221	-2.257	1.398	.026*
	† $\log_e$(Time on site/app)	-0.129	[-0.287, 0.028]	0.080	-1.624	1.208	.107
	Z-score Compulsive Use	0.171	[-0.025, 0.362]	0.098	1.719	1.208	.088
3	Constant (β_0)	0.001	[-0.319, 0.322]	0.162	0.009		.993
	Strict diet	-0.989	[-1.427, -0.551]	0.221	-4.471	1.409	.000***
	Self-regulated diet	-0.509	[-0.948, -0.070]	0.222	-2.293	1.403	.024*
	† $\log_e$(Time on site/app)	-0.129	[-0.288, 0.029]	0.080	-1.619	1.211	.108
	Z-score Compulsive Use	0.073	[-0.292, 0.439]	0.185	0.398	4.260	.691
	Strict diet x Compulsive Use	0.262	[-0.245, 0.769]	0.256	1.024	2.003	.308
	Self-regulated diet x Compulsive Use	0.060	[-0.376, 0.496]	0.220	0.273	3.098	.785

* p < 0.05, ** p < 0.005, *** p < 0.001
Model 1: $\log_e(\frac{y_1}{y_0}) = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + e$, (Adj. R^2 = 0.126)
Model 2: $\log_e(\frac{y_1}{y_0}) = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_3 \log_e(x_1) + \beta_4 x_2 + e$, (Adj. R^2 = 0.1396)
Model 3: $\log_e(\frac{y_1}{y_0}) = \beta_0 + \beta_1 d_1 + \beta_2 d_2 + \beta_3 \log_e(x_1) + \beta_4 x_2 + \beta_5 d_1 x_2 + \beta_6 d_2 x_2 + e$, (Adj. R^2 = 0.134)
† This variable refers to the baseline period, to aid interpretation it has been centered $(x_i - \bar{x})$

The results of the first model in Table 4.3 show that the strict newsfeed diet is a significant predictor of log change ($p < 0.005$) whereas the self-regulated diet intervention is a suggestive predictor ($p < 0.05$). These predicted changes mirror the descriptive results seen in Table 4.2 and raincloud plot D) in Figure 4.3, while controlling for the effects of the other intervention. The first model suggests that if a user was in the strict newsfeed diet group (while controlling for the effect of the self-regulated newsfeed diet) their change in time on Facebook's site/app between the baseline and treatment periods was $(\exp(\beta_1) - 1) * 100 \approx -63.868\%$.[‡] H2a is sup-

[‡]These percentage changes in time on site/app refer to the natural log transformed space. As a

ported. If a user was in the self-regulated diet group (controlling for the effects of the other group) their change in time on site/app between the baseline and treatment periods will be approximately $(\exp(\beta_2) - 1) * 100 \approx -40.310\%$. ‡ H2b has suggestive support. The results of the second model indicate that the interventions remain stable predictors of log changes in time spent on the platform when controlling for the users' logged baseline measures for time on the site/app and their standardized compulsive use before the baseline period. The increase in explained variance (Adj. R^2) between the first and second models is non-significant: R^2-change .026, F-change $(2, 129) = 1.983, p = 0.142$. That is, users' time on site/app during the baseline period and their level of compulsive use only explains an additional 2.6% of changes in time on the platform when we have already controlled for the effects of the two newsfeed diets. Also, the second model shows that neither the users' time on site/app during the baseline period nor their level of compulsive use are suggestive predictors of log changes in time on site/app, see Table 4.3. So while compulsive use has a strong association to average daily time on site/app during the baseline period, it is not a strong predictor of *changes* in time on site/app between the baseline and treatment periods when controlling for the other predictors. Lastly, the results of the third model assess if the log changes in time on site/app for each standard deviation increase in compulsive use is significantly different between users who are in the intervention groups (while controlling for the previously mentioned effects in the first and second). As Table 4.3 shows, the moderating effects of compulsive use on the strict and the self-regulated newsfeed diet add basically no explanatory value to the third model: R^2-change .008, F-change $(2, 127) = 0.603, p = 0.549$. The regression coefficients for the interaction terms $\hat{\beta}_5 = 0.262$, $p = 0.308$, $\hat{\beta}_6 = 0.060, p = 0.785$ respectively, are not suggestive

reference point, the untransformed median percentage changes in time on site/app between the baseline and treatment weeks for each newsfeed diet are approximately: No diet: -6%; Strict diet: -70%; Self-regulated diet: -45%.

predictors of change in time on site/app. In fact, the proportion of explained variance goes down between the second and third models. These moderation effects (or lack thereof) suggest that the efficacy of the two newsfeed diets are not impacted by more and less compulsive users. H3 is not supported.

Effects of restricting the newsfeed on user experience (H4a–H4b)

To investigate H4a and H4b we coded and analyzed the answers for the SUS questionnaire and the open usability questions, for which 140 users provided answers (367 comments, 6608 words). The SUS score, which focuses on the ease of use of Facebook did not change significantly between the baseline and the treatment periods. The answers to the open questions were coded line by line by two of the co-authors based on the negativity of the emotions related to the usability, such as negative sentiment, hate, frustration, boredom or annoyance about missing important information. For instance, the following answer was coded as negative: "*It was annoying not getting updates from friends, but particularly from my groups – I was actually trying to sell and give away some items in some groups and I had to find the notifications and information and messages manually!*"

To validate the codes, we measured inter-rater reliability [290]. The two coders agreed on 135 ratings and disagreed on 5. That translates to 96.4% agreement which is above 75% and hence considered acceptable [291]. The coders discussed the disagreements to reach a consensus. The results show a proportion of negative comments of 23.3% in the control group, 21.2% in the self-regulated diet group and 56% in the strict diet group. The chi-square test of independence showed that the proportion of users reporting negative comments differed significantly between the newsfeed interventions $\chi^2(2, n = 140) = 16.316$, $p < 0.001$. Additional χ^2tests showed that users in the strict newsfeed diet condition reported a significantly more negative user experience than the control condition $\chi^2(1, n = 93) = 10.258$, $p < 0.001$ and users in the self-regulated diet group reported significantly less negative user experience than the strict newsfeed diet condition $\chi^2(1, n = 97) = 12.259$, $p < 0.001$. These results are supportive of H4a and H4b.

The content of the negative comments was different between newsfeed diet groups. In the control group, with no diet, some users had the impression that the extension that they installed affected their user experience of Facebook, even though the extension was not active in this condition. Consequently, most negative comments in this condition mentioned unspecific effects such as: "*Your feed can get a bit boring*", "*Feeling anxious about maybe missing out on something important*", "*Feeling like you missed out*", "*When I opened Facebook it would show a lot of white at first. It worried me what I would miss.*" In the strict newsfeed diet, in contrast, the negative comments mentioned specific effects such as "*It was annoying that there was absolutely nothing left on my Facebook feed*", "*I actually missed out on some key things referenced by friends in conversation mentioned on Facebook*", "*It unfollowed some people and caused me to miss some important posts.*" The negative comments of the self-regulated diet condition were overall similar in tone to the strict newsfeed diet condition such as: "*I lost everyone who I enjoyed following*" or "*I wasn't able to keep up with any of my family or friends that live far away.*"

4.5 Exploratory Thematic Analysis

To go beyond the results related to the hypotheses from Section 4.2 and to broaden the analysis of the user experience, we conducted an exploratory thematic analysis [292, 293] of the three open usability questions about their positive, negative and overall user experience. We first established reliable codes for user experience and then used those codes to generate themes representing shared meaning. We used an inductive and deductive hybrid approach [294]. Deductive codes were developed such as disconnect, mindfulness and autonomy, to name a few based on previous research on user experience that we reviewed in Subsection 2. Inductive codes were added after reviewing the data. Then we iteratively coded the 140 responses over a period of one week. During our coding, we focused on responses indicating user experience. We excluded several responses that did not relate to user experience but rather referred to some technical glitches during the experiment that participants experienced such as, "*I want to say that I messed up when creating a shortcut when I close Facebook on my iPhone. I rechecked my mistake and saw that the file path was the same as /Open.csv. I am sorry for my mistake.*" Also, the responses such as "*study was fine*" and "*I think it was well done*" were also not taken into account. The process of coding was done by two researchers independently. The researchers then came together to share their independent analysis, which was discussed, and relevant themes were generated. In total, six main themes emerged and are discussed in the following section: Fear of missing out (FOMO), focus, self-awareness, ease-of-use, sense of control and liberation. These results are summarized in Table 4.4 on the last page of the chapter.

Fear of missing out (FOMO). FOMO is having a persistent fear that others may be having rewarding experiences while one is deprived [295]. More precisely, it is people's fear of missing out on experiences across their extended social network. The users in the field experiment experienced the fear of missing out despite knowing that they could contact their friends via messages on Facebook. This suggests that these relationships can be quite passive or dormant on Facebook. The responses such as "*Feeling anxious about maybe missing out on something important*" were coded as "*missing*". One user reported that they actually missed out on an important event: "*I missed my friend's memorial service because I did not see the posts about her death.*" It also came to light that participants used Facebook to track the wellbeing of their friends and family: "*I would miss out on important things that are happening with friends and family.*" In the strict newsfeed diet a participant mentioned, "*But I did find myself wondering quite often what I might be missing. I worried particularly about missing some important news in my friends' lives.*" Not just the feeling of missing out but the feeling of being judged by others also emerged: "*I was worried people thought I was ignoring them or didn't care about what was happening in their lives.*" Figure 4.4 suggests that the members of the self-regulated diet were less likely to worry about missing out than those in the strict group, yet they reported that they received fewer posts on their newsfeed from their friends and family than they had expected: "*It doesn't update with new things that you've followed often enough.*" In Figure 4.4, the FOMO theme emerged about twice as often for users in the strict diet compared to the other newsfeed diets. Figure 4.4 also highlights that there is a suggestive association between the occurrence of this theme and the newsfeed interventions (χ^2 = 6.356, V = 0.213, p = 0.042). As an interesting side note, some users in the control group reported seeing fewer posts after installing the chrome extension, despite no changes on their newsfeed being made by the extension.

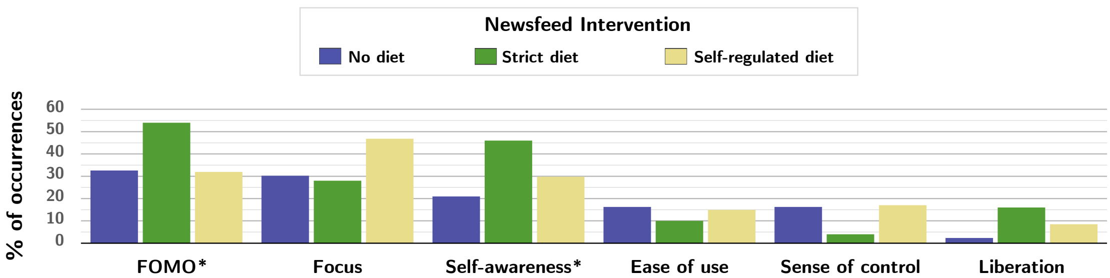

Figure 4.4: Percentage of users in a given newsfeed intervention group who mentioned a given theme. χ^2 tests were conducted to test the likelihood of occurrence of a given theme across the interventions $(2, n = 140)$, * p<0.05, ** p<0.005, *** p<0.001. These tests were performed without using a family-wise error correction, i.e., $(\alpha_E/1)$.

Focus. In our analysis we consider focus to be a state when individuals are able to direct their attention to meaningful activities without being distracted. Previous studies suggest that users are often preoccupied with distractions, which leads to difficulties focusing their attention on other tasks [296, 266]. The responses such as "*my Facebook experience became more wholesome overall*" were coded "*UX improved*". While the responses that referred to the improvement of their ability to focus, such as "*it cleared my Facebook newsfeed of post and pictures I didn't want to see, which I really liked. Ever since downloading the extension I see more relevant and important things on Facebook*" were coded as "*declutter*".

Our analysis indicates that individuals were able to focus on tasks after cleaning their newsfeed. For instance, some individuals felt that the newsfeed intervention helped them to focus by removing distractions, while keeping relevant content on their newsfeed, (e.g., "*The best thing is that it can help reset your Facebook feed, and by that I mean that it can help clear your Facebook feed so that you can go back and choose what you want to see everyday. It can help declutter your feed"*). A user from the self-regulated diet reported an increased ability to focus on other tasks: "*there was more time to complete other tasks.*" Likewise, another user reported that the unfollow process has enabled focusing on spending more time with family members: *"I felt slightly disconnected from others but it was kind of good because I spent more time with my family than scrolling Facebook."* Not only did users report reduction in distractions but they indicated that the posts on Facebook newsfeed had become more relevant, allowing them to focus on relevant posts: "*ever since downloading the extension I see more relevant and important things on Facebook."* Figure 4.4 highlights that users in the self-regulated diet condition reported often that cleaning the newsfeed has led to a decreased number of notifications that usually leads to distraction: "*the best part is not being overwhelmed with several notifications*". Figure 4.4 describes a more frequent occurrence of "*focus*" as a theme in the self-regulated newsfeed compared to the other two groups. However, there was no suggestive association between the occurrence of this theme and the different newsfeed diets ($\chi^2 = 4.387$, $V = 0.177$, $p = 0.112$).

Self-awareness. Our study referred to the following definition of self-awareness: the ability to reflect and evaluate oneself objectively through introspection [297]. Most often, individuals open the app unconsciously after a certain period of time, eventually leading to automatic behaviors of checking social media regularly and not being mindful of their habit [298, 299]. When participants reported several responses such as the following "*in a way helped me to see exactly how much time I inadvertently spent on Facebook and how often I went on there*" we coded them as "increased consciousness." Some comments suggested that the participants began to get out of the habit loop of opening Facebook often, especially the users in the strict diet group: "*After a while I just got out of the habit of checking Facebook on my laptop*". Moreover, many participants became self-aware that spending time on Facebook was not the best allocation of their time: "*Not seeing my feed on Facebook for 2 weeks convinced me that I really don't need Facebook anymore and that's a good thing*". Also, the reason behind the participants' increased Facebook use came to light as individuals' Facebook feed became more empty and decluttered: "*Interesting to see how often I look to Facebook to relieve boredom.*" The participants in the self-regulated diet group often reported that they decreased the time they spent on Facebook: "*it caused me to spend less time on Facebook*" and the intervention made them realize that they do not need Facebook: "*I found that I don't need Facebook overall, and this study helped me figure that out.*" While the only difference between self-regulated and strict diets was that participants in the strict diet were forbidden from following their friends, pages and groups, the theme of self-awareness emerged more often among those on the strict newsfeed diet compared to the other diets, as shown in Figure 4.4. There is also a suggestive association between the occurrence of this theme and the newsfeed interventions (χ^2 = 6.888, V = 0.222, $p = 0.032$).

Ease-of-use. Previous research found that platforms with fewer frictions or interruptions and greater ease of use improve user experience [300]. Our analysis followed previous research and deemed reduced effort an indicator of ease of use. Users generally did not need to put a lot of effort into unfollowing their contacts. Several participants emphasized "*ease-of-use*" which is an essential component of usability. The responses such as "*I appreciate how the extension has the ability to automatically unfollow friends, pages, and groups*" were as coded as "*effortless*." The users appreciated the ease-of-use of the intervention as it worked in the background, and individuals did not have to manually unfollow any of the friends, pages and groups: "*I like that it automatically unfollowed pages for me since I followed a lot of them.*" Furthermore, users appreciated that extension does not interfere with their other Facebook activities: "*It is very hidden and does not interfere with my actions.*" A code that kept recurring in the strict diet condition was automated: "*The best thing about the extension is that it does everything for you automatically.*" The difference between the interventions regarding how many times their responses came under the theme related to ease-of-use is non-significant, as seen in Figure 4.4, (χ^2 = 0.883, V = 0.079, $p = 0.643$).

Sense of control. Sense of control is a psychological phenomenon that occurs when individuals feel that their actions and consequences are under their control, i.e., they feel that they are in the driving seat [301]. The newsfeed diets seemed to instill in users a feeling of autonomy and control. The idea of automating the unfollowing of contacts that users have been following by default and then allowing them to choose which friends, pages and groups to follow increased their sense of agency and control on Facebook: "*the best thing is that it can help reset your Facebook feed, and by that I mean that it can help clear your Facebook feed so that you can go back and choose what you want to see everyday.*" Another user from the self-regulated diet reported that, "*It allows me to have control over my Facebook account. I enjoy the overall user friendly interface and design.*" Responses of this kind indicate that when the sense of agency increases, the user experience may be enhanced. One user mentioned the following: "*It allowed me to selectively follow people I really care about. When choosing people to re-follow I found there are a lot of people I don't really care to follow. I also didn't feel guilty because I knew that they wouldn't know I wasn't following them.*" It is important to mention that Facebook's UI already gives users the option to selectively follow contacts they really care about. As we mentioned in the section on related work, this option exists on Facebook, but can be difficult for users to discover.

In Figure 4.4, we also observe that users in the strict diet condition did not feel as much of a sense of agency or control as those in the self-regulated condition. That said, there was no suggestive association between the newsfeed diets and the occurrence of this particular theme ($\chi^2 = 4.846$, $V = 0.186$, $p = 0.089$).

Liberation. When individuals are able to put an end to their dependence on a certain activity or pattern, they feel liberated [302]. With regard to Facebook use, liberation is when individuals lose their desire or motivation to use Facebook, which they had been captivated by. Several users reported a sense of liberation as their motivation to use Facebook declined over the course of the field experiment. Few users in the self-regulated diet reported a loss of motivation and desire to use Facebook, see Figure 4.4.

The users in the strict diet condition reported more frequently that their motivation and desire had taken a hit: "*I am less motivated to use Facebook when I should be* *doing something else, like working or studying.*" As one user reported, the lack of motivation and desire also led them to question their need for the service itself: "*Not seeing my feed on Facebook for 2 weeks convinced me that I really don't need Facebook anymore.*" Sometimes users in the strict diet reported that the resulting friction around the access to posts and information as something positive: "*The best thing about the app is it really removed my desire to go onto Facebook. I couldn't easily see posts or information from any of my friends, so I didn't feel very motivated to check the site at all.*" However, we saw no suggestive association between the newsfeed diets and the occurrence of liberation as a theme (χ^2 = 5.182, F = 5.012, $p = 0.076$)†.

4.6 Discussion

The present chapter investigated the effects of a novel automated digital unfollow intervention that restricted and practically reduced the content of individuals' newsfeed by unfollowing the users' friends, pages and groups. We conducted a field experiment with real Facebook users randomly assigned to one of three independent conditions: the control condition (no diet), a strict newsfeed diet and a self-regulated newsfeed diet. Unlike other digital wellbeing research that focuses on external interventions such as timers [272], overlays, feedback [102] and limits [303], we focused on changing the internal choice architecture of the platform itself. We structure the following three subsections by laying out first our contribution, and we close the subsections with the implications to the relevant stakeholders.

†F refers to Fisher–Freeman–Halton's Exact test statistic as the cell count for this particular theme is < 5 for the No Diet intervention

4.6.1 Newsfeed diet interventions

Our results show that users reduce their Facebook use just due to the restriction or reduction of their newsfeed. We observe that a strict newsfeed diet is a significant predictor of change in time on site/app (H2a) by around -64%. Understanding exactly how individual design mechanisms translate into more or less usage of a social media app is difficult owing to the subtle nuances in the design space of previous interventions. The results of a previous study that related to the Facebook accounts of several US police departments indicate that posts containing UI elements such as links and images translate into more likes and interaction from users [304]. This suggests that if such UI elements are limited upstream (e.g., by a newsfeed diet) then users would spend less time on the platform itself. However, previous research results related to filtering or blocking the newsfeed are mixed. As we mentioned in the section on related work, Lyngs et al. conducted a user experiment where one of the interventions was a completely hidden/blocked newsfeed [266]. This newsfeed blocking intervention did decrease users' visit length on Facebook's site with FOMO being a major theme among the users. While we did not use visit length as a dependent measure, our overall results are quite well in line with those findings. Another study found no significant differences in usage time when introducing a feed filter on Twitter [276]. However, this feature was partially responsible for increasing users' sense of agency on the platform itself. In the strict condition, we observed that the drop in time on site/app was happened in parallel to more frequent reporting of negative user experiences, as measured by negative emotions (H4a). Yet, our thematic analysis provides a more mixed picture of this result where users in the strict condition more frequently express rather positive themes such as liberation and self-awareness. These results indicate that such a drastic intervention may be effective at reducing time spent on Facebook, but is not satisfactory from a user perspective. This result is also in line with previous studies that have underlined that more restrictive mechanisms cause more frustration among the users [305, 266]. Previous design researchers have attributed this reaction to the diversity of usage contexts and user needs [305, 269].

In addition to these observations, we also found that the self-regulated newsfeed diet had a suggestive decrease in time on site/app (H2b) by around 40%. Furthermore, our findings indicate that this type of intervention is associated with a more positive user experience as compared to the strict newsfeed diet condition (H4b). Indeed, over half of the comments in the strict diet condition reported negative feelings compared to only around one in five in the control and the self-regulated diet conditions. This is a novel contribution to the field, which indicates that not only complete abstinence but also limitation of the number of content providers of the newsfeed may reduce time spent on the platform while enhancing users' experience compared to more strict abstinence. Looking in more details at the user experience, our thematic analysis provides a first tentative explanation that such a self-regulated diet helps users to avoid distractions and to focus on the useful interactions that a social media newsfeed can provide.

Future research could further explore the mechanisms through which modifications in the newsfeed affect user experience and time on site. For instance, our research cannot answer questions about whether the observed effect is due to the number of connections unfollowed, the type of connections unfollowed or the type of content that appeared in the newsfeed itself. In a nutshell, these questions relate to the social graph architecture which connects nodes (users), through edges (follow relationships), on which messages (content) can be passed. This leads to future research questions such as how do social graph architecture components affect time on site and user experience?

Also, future research could further investigate the value provided to users by different types of connections. Our thematic analysis related to FOMO provides some preliminary indications that users valued interaction with friends and family, which aligns with some previous findings [259, 179]. Future research could, however, formally confirm whether that value enhances user experience.

Implication for designers

Our results suggest that the current newsfeed design might work as expected by social media platforms: it increases time spent on site compared to more curated versions. In this regard, social media platforms have little direct incentive to give users the tools to easily unfollow those connections that are less interesting to them. However, approaches that increase user benefits at the expense of time spent on the platform (resulting in potential short-term revenue decline from advertising) could give longer-term benefits to the platform, such as increased loyalty, increased brand image and less regulatory scrutiny. Furthermore, potential FOMO associated with cutting ties with updates from friends and family might indicate that social media companies could build on these connections rather than on unconnected content updates, which seems to be a current trend. Recent reports suggest that Facebook will modify the platform's newsfeed feature to focus more on "*unconnected*" content sources in an effort to compete with TikTok [306, 307]. Finally, from the perspective of designers of apps, those plug-ins or extensions that help users spend less time on social media or at least help them regain control over their time on social media may target the newsfeed to effectively intervene in the platform. Removing the newsfeed or limiting its content are useful tools in that respect.

4.6.2 Understanding compulsive use

In our research, and contrary to what we predicted, different levels of compulsive use did not moderate the effectiveness of our interventions (H3). These results are surprising, since they contradict what could have been predicted based on previous literature on habits [273]. Nevertheless, the results seem to confirm more recent literature on closely related digital wellbeing interventions [269]. This is an interesting prospect, showing that digital interventions can curb Facebook use even for compulsive users. Our thematic analysis hints at potential novel mechanisms explaining the impact of that type of digital interventions on time on site. For instance, the strict diet intervention could increase users' self-awareness, which allows them to reflect on their behavior as well as to decrease the motivation to go online due to a potentially less rich user experience. According to our data, these mechanisms would be operating irrespective of the user's level of compulsive use.

In addition, previous scholars have argued for more comprehensive research into the underlying mechanisms that may influence users' compulsive behavior on social media platforms [245]. In this respect, our study contributes to quantifying the antecedent of users' compulsive use. We have investigated the association between the time spent on Facebook's site/app and compulsive use (H1). Our data suggests a clear statistical relationship between time on Facebook and self-rated compulsive use. However, the Pearson correlation between compulsive use and log transformed time on site/app results in a coefficient of determination of $(0.388^2 = 0.151)$. That is, only about 15% of the variation in compulsive use is explained by its relationship with time on site/app. This means that roughly 85% is still unexplained. In other words, there are several other relevant factors that are associated with compulsive use.

Future research could allocate more attention to the effect of newsfeed (and more broadly digital) interventions on compulsive users. One such research direction could further dissect time spent on site/app as the measure itself might not be enough to account for digital wellbeing [276]. The type of interaction and the context of the interaction could be further explored. Also, it could be useful to look in more details at cases at the margin, such as high use but low compulsion or high compulsion but low use, which could shed further light on how to best design social media-like interactions in the interest of users. Furthermore, although we show that 15% of the variation in compulsive use is explained by its relationship with time on site/app, we do not know how a decrease in time on site/app affects the level of compulsiveness.

Implication for designers

The major implication is that newsfeed modifications can significantly affect the time on site/app of compulsive users and, as a result, serve as a potential tool to increase digital wellbeing. Furthermore, policy designers who are currently interested in curbing dark patterns online [308] could leverage our results to push for better support for compulsive users and push social media companies for more accountability on this matter.

4.6.3 Designing for agency

Previous HCI researchers have demonstrated the importance of users feeling in control of their interaction with social media platforms [259, 266, 276]. As described in cognitive neuroscience, control refers to being in charge of one's own actions and affecting the external environment through that control [309]. Through years of scientific research and testing, it has been demonstrated that a positive user experience is associated with a sense of agency [310]. In this vein, a recent study showed that some users are looking for some kind of middle ground, such as newsfeed featuring only their best friends or favorite pages [179]. Another study by Lyngs et al., [266] found that most participants expressed a desire for easy ways to filter, limit or disable newsfeeds.

Against this backdrop, our study provides a novel intervention to reduce use while giving users more control over their newsfeed. Furthermore, the thematic analysis highlights that users in the self-regulated diet as well as in the no diet (control) condition, reported a sense of control that was mostly missing from users in the strict diet condition. As a result, while both interventions led to less time spent on Facebook, only the self-regulated diet reported a retained a sense of control that was similar to the no diet group. It should be noted that with current mainstream social media platforms, users could potentially implement such a self-regulated diet. However, as mentioned above, the default option which follows all contacts and forces users to explicitly opt-out works in a way that agency is reduced [64]. Future work could further investigate restrictions of the newsfeed while maintaining user agency. For instance, reversing the default in terms of following connections, from opting out to opting in, may give users more control. This would look like our self-regulated diet intervention, except that users would not have been exposed to the control (i.e., everyone followed by default) beforehand. Another possibility could be to follow by default close connections (e.g., start following users with whom there are messenger interactions), but not other types of connections. If needed, users could then follow other connections by themselves.

Implication for designers

Our findings suggest to designers of apps, plug-ins or extensions that whereas a strict diet feature might be more powerful, if followed by the users, it may trigger more reactance because of the freedom of restriction that it entails. In this case, users might more easily abandon it than they would with a self-regulated diet. Facebook and social media platforms seem to have limited incentives to give users agency in a way that would reduce their time on site. However, policymakers (e.g., regulators) could give users more agency by ensuring that they have the possibility to implement their intentions on the platforms that they are using. This is what the GDPR offers with the prohibition of the opt-out policy. Closer to our context, the regulator could enforce a principle of fair and symmetric frictions in user choice. If following all friends on a platform like Facebook is easy and friction-less (because it is the default option), unfollowing all friends could be made as easy as one click, in order to let users re-follow just their few important connections. Such a design principle could have broader reach than just the newsfeed. For instance, if creating an account needs two clicks, regulation could force the platform to allow the removal of users' account in two clicks.

4.6.4 Limitations

This study faces several limitations that could be addressed by future research. First, the Chrome extension's unfollowing procedure could take a long time depending on how many contacts users have. Consequently, the results suggest that some users in the strict unfollow condition did not fully unfollow all contacts as planned (the median percentage drop in followed contacts was 82% for this newsfeed diet). As such not all users in the strict diet had an empty newsfeed. Also, a given user's newsfeed could still potentially contain sponsored content even if all contacts are unfollowed. However, in the strict diet condition, even if the newsfeed was not empty for all users, no user could re-follow a contact in the browser. In this way, the fundamental difference in control over the newsfeed across conditions remained. Second, our data did not allow us to infer precise refollow behaviors. Identifying users' behaviors and attitudes behind the use of the newsfeed and further exploring who and why they chose to re-follow specific friends, pages and groups could have contributed to insights that would have been very helpful when designing for digital wellbeing in social media contexts. Third, our study only measured time on site/app on one particular social media platform, namely Facebook. We can therefore not assess whether the changes in time spent on Facebook was replaced by more time spent on other platforms such as Instagram or TikTok.

4.7 Conclusion

In this chapter, we tackled the pressing issue of compulsive social media use. We took a design science approach to design a solution to reduce Facebook use specifically. Our solution – which consisted of reducing the endless newsfeed, breaking the relationship, called following, which makes a contact's updates spill over into the user's newsfeed – does indeed reduce the time spent on the platform, but it can have negative consequences in terms of usability. An alternative approach that lets users self-regulate their newsfeed diet suggests a decrease in time on the platform – without steep costs to the user experience. Our results also indicated that both of these approaches seemed to work for more compulsive as well as for less compulsive users.

Table 4.4: Main themes and their associated codes from the thematic analysis with representative quotes from each newsfeed intervention.

Main themes	Associated codes	Representative quotes related to each newsfeed intervention
FOMO	missing, anxious, missed information, fewer connections, less content, disconnection	*No diet:* “Feeling like you missed out” *Strict diet:* “I was not updated about anything on Facebook” *Self reg. diet:* “Not being able to see a lot of content on Facebook, only seeing the same content over and over”
Focus	declutter, less notifications, reduced distraction, time for other tasks, UX improved	*No diet:* “Seeing fewer notifications” *Strict diet:* “Focus on my work and accomplish what i needed to” *Self reg. diet:* “It allowed me to selectively follow and focus on posts that I really care about. When choosing people to re-follow I found there are a lot of people I don’t really care to follow. I also didn’t feel guilty because I knew that they wouldn’t know I wasn’t following them”
Self-awareness	less time-spent, behavior change, increased consciousness, mindful	*No diet:* “The mental clarity it provides” *Strict diet:* “In a way helped me to see exactly how much time I inadvertently spent on Facebook and how often I went on there” *Self reg. diet:* “I have decided to reduce my time using social media. Also, I removed all of them from my phone 2 days ago. This survey probably led to big changes in my life and I hope it will change my life for better”
Ease-of-use	automation, effortless, painless, automatic	*No diet:* “I was expecting it to feel a lot more intrusive but it wasn’t” *Strict diet:* “The best thing about the extension is that it does everything for you automatically” *Self reg. diet:* “It is very hidden and does not interfere with my actions”
Sense of control	manage newsfeed, ability to control, capability, autonomy	*No diet:* “It gives you more power to control what you see and don’t see on your Facebook feed, which could make your experience more positive. Things that irritate you will not irritate you anymore” *Strict diet:* “Now I can choose what I actually care about to pay attention to” *Self reg. diet:* “It makes you deliberately decide who to re-follow and what fits your interest”
Liberation	stopped using Facebook, reduced desire, reduced motivation	*No diet:* “It more or less will defeat the purpose of continuing to have a Facebook account. You might as well either not log on or deactivate your account” *Strict diet:* “Not seeing my feed on Facebook for 2 weeks convinced me that I really don’t need Facebook anymore” *Self reg. diet:* “There’s no point in checking it”

5 Stacking Nudges for Online Charitable Giving

This chapter is based on a working draft paper: **Kristoffer Bergram**, Abdessalam Ouaazki, Manon Berney, Valéry Bezençon, & Adrian Holzer. *Stacking Nudges for Online Charitable Giving: An exploratory analysis of pro-social nudge combinations and their polarizing effects from a user's perspective.* A working paper, University of Neuchâtel.

Contents

5.1 Introduction

Since 1950, charitable giving has steadily accounted for about 2% of the American economy [311]. In 2022 alone, Americans gave away roughly 500 billion dollars to charity and about 64% of that came from private individuals [312]. For several years, online giving has increased by about 10% on a year-on-year basis [313, 314]. Online channels are also progressively used for various kinds of charity donations across the globe in countries like India, China, Kenya, and Brazil [315]. It is estimated that even before the Covid-19 pandemic, more than 60% of non-governmental organizations (NGOs) worldwide accepted online donations on their websites [316]. In the same survey, respondents from more than 5700 NGOs, rated their websites as the most effective tool for fundraising [316].

The act of donating to charities is increasingly performed through digital channels. This means that someone (usually some kind of web or app designer) has to aid an NGO in setting up a digital choice architecture for online users that maps various design elements to different donation options. This kind of designer can also be called a choice architect [1]. Charities will inevitably need to decide on a digital choice architecture around how they acquire online donations from users, and these design decisions will likely impact users' downstream donations [317, 31]. That is, online donors will inevitably be nudged in one or another direction. In the context of charitable giving, digital nudges will therefore play an increasing role. Digital nudging entails using UI design elements that guide people's choices or behaviors in digital choice environments [27, p. 1].

Large charities already deploy various kinds of nudges in their effort to elicit donations online [317]. Nudges in pro-social domains such as charity are unique in that they are often aimed at steering people's behavior in the direction of some kind of charitable or public welfare, rather than private welfare [318]. As digital nudges become more pervasive in several application domains, users become more likely to be exposed to several nudges at the same time, the effects of which require more research [147, 30, 319]. While previous studies have investigated pro-social nudging in the charity donation context (for example [320], [321], [322], [323], [46]), offering significant contributions to several facets of pro-social nudging, they rarely investigate the systematic effects of stacking or combining several nudges at once. Several recent studies have investigated combinations of digital nudges, but these studies have been conducted across diverse application domains such as e-commerce [105], health [118], energy-saving behaviors [324], privacy/security [319, 325], and social media [88]. Because of the diversity of both their design elements and their application contexts – it is challenging to paint a comprehensive picture of how various nudge combinations relate to users' online donation behaviors. Online charitable giving is a special domain in the sense that there is an inherent trade-off between the clear altruistic act of donating to charities vs. just keeping our monetary resources for ourselves [322]. To the best of our knowledge, no study has so far investigated interactions between digital nudges in this pro-social context. Stacking several nudges intuitively entails putting more pressure on users to act in a particular way [324]. Also, backfiring effects have already been documented in the context of online fundraising related to the frequency of reminder emails [326]. This means that online charities may have to perform a difficult balancing act when it comes to steering their potential donors in a more benevolent direction.

In this exploratory chapter, we will use three common digital nudge patterns [30, 73], and investigate the potential interactions of their combinations related to users' donation behaviors. We perform this evaluation in a pro-social context with real monetary stakes i.e. where users give away a real portion of their earnings. We will also explore several ethical dimensions of these pro-social nudge combinations from the user's perspective. In a recent review article, Clausen et al. [327] highlight that additional outcome measures (apart from the immediate behavioral outcome) such as various cognitive or emotional strains on the user should be considered when designing persuasive technologies. While digital nudges have attracted more and more attention from researchers in recent years, Susser and Grimaldi [328] point out that existing empirical research investigating the effectiveness digital nudges and targeted recommendations tends to neglect other ethically relevant effects on users. This raises relevant empirical questions around users' perspectives on how or why they are being nudged in particular directions [328]. Given the unique nature of pro-social nudges (that they guide the user's behavior in the direction of public rather than private welfare [318]), we will also explore the relationship between these nudge combinations and two important ethical concerns associated with nudging [317, 50], namely perceived threats to autonomy and sense of manipulation (from the users' perspectives).

5.2 Methodology and Approach

This chapter will investigate the intersection of digital nudge combinations and online donations by using a design science research methodology (DSRM) approach [329, 212, 251]. The rest of this chapter is structured according to the six steps of DSRM as outlined by Peffers et al. [212]. The previous section described the boundaries of the problem that we address. Section three further specifies the objectives of our solution in the context of prior literature. Section 5.4 covers our design decisions and section 5.5 is a demonstration of a charity donation UI for testing combinations of digital nudges. Section 5.6 addresses the exploratory research questions by evaluating these nudge combinations in relation to users' donation behaviors and users' self-reported experiences around autonomy and manipulation. The final section surrounds a discussion of our findings.

5.3 Background and Related Work

We will now outline previous work that has investigated different digital nudge combinations, the relationship between nudging, autonomy/manipulation, and the potential knowledge gaps that remain related to these areas.

5.3.1 Combining digital nudges

Online users are increasingly likely to be exposed to several digital nudges simultaneously and several previous scholars have called for more research that investigates the potential combinations of digital nudges [147, 30, 319, 330]. Several recent examples of nudge interaction studies ranging from application contexts privacy [319], health [118], or e–commerce and marketing [105] can be found in the literature.

However, these interaction studies have shown mixed results in different application contexts. There is evidence suggesting that stacking several distinct nudges on top of each other yields a larger effect in terms of behavior change across several application domains (see for example [324], [105], [86], [111]). For example, in the context of online grocery shopping, scholars have investigated combinations of defaults and social nudges and shown that when combined, they can lead to a stronger impact on compliance with product recommendations than each nudge individually [105]. Kretzer and Maedche [86] have shown that a variety of social nudges can moderate each other in a BIS context. They relied on a feature for report recommendations (that employees were recommended to read through an adjacent link) that utilized various social cues for the recommendation itself. Examples of social nudges included institutional proximity to the user that provided the report recommendation and social hierarchy i.e. whether the person who recommended the report was an intern, manager or director. Their results indicated that social nudges in this context moderated each other's efficacy. One example was that the acceptance rate for reports from interns was increased with institutional proximity, but this was not true for report recommendations from directors [86]. In the context of crowdfunding, researchers have revealed significant interactions between social nudges and the amount of discount present for various backer options [111]. These two design cues, discounts and the number of backers for a given reward option are common features on crowdfunding platforms, and the effect of a given discount seems to be moderated by the popularity of a given project in this application domain [111]. In the context of online privacy, researchers have investigated two-way and three-way combinations of default, friction and social nudges on user consent screens [319]. Shore and Cummings [319] investigated how the above nudges influenced users' tendencies to allow the disclosure their location data in an online experiment. While they did not uncover a clear interaction pattern between these digital nudges, they did conclude that the effect of friction seemed to be dependent on the presence of other design elements [319]. In another recent study aimed at steering users towards greener fashion products in an e-commerce scenario, the results suggested a backfiring effect from employing a combination of digital nudges [330]. This pertained to default and social nudges specifically. Another study on energy-saving behaviors among employees suggests that using single nudges had no effect. In contrast, stacking three different nudges on top of each other had complementary effects in terms of pro-environmental behaviors [324].

Findings like these indicate potentially relevant interaction effects for a variety of digital nudges across several application domains. Yet, taken together, previous studies do not form a conclusive picture of when certain digital nudges interact, or if these effects can be extrapolated to pro-social contexts such as online charitable donations. Our first objective is to investigate this intersection of digital nudge combinations and online donation behaviors. This leads to our first exploratory research question:

RQ1: How do pro-social nudge combinations affect users' charity donations?

5.3.2 Autonomy and manipulation as ethical concerns

In broad terms, nudging entails interventions that steer people in a particular direction while still allowing them to go their own way [31]. Thaler and Sunstein [1] underscored that when nudging was applied, it should make people better off, as judged by themselves. The general discussion around the ethics of nudging often boils down to two principal concerns: autonomy and manipulation [50, 317]. Autonomy as an ethical consideration is nuanced and has often been conceptualized in diverse ways across the nudging literature [331]. As an ethical concept, autonomy pertains to the control people have over their own evaluations and choices [63]. In their review on nudging and its relationship to autonomy, Vugts et al. [331] clarify three main strands of autonomy that are frequently discussed in the nudging literature: freedom of choice, agency, and self-constitution. We will mainly investigate autonomy as it relates to freedom of choice. While the label "manipulation" has quite pejorative connotations [63], we have adopted it here since it is a widely discussed ethical concern in relation to nudging (see for example [56], [332], [333] for more comprehensive discussions of this topic). Sunstein [31, p. 443] demarcates an action as manipulative when it attempts to "... influence people in a way that does not sufficiently engage or appeal to their capacities for reflective and deliberative choice". The point has also been raised that manipulation often has the distinguishing mark of a justified sense of ex post betrayal [334]. This means that it may not be until after the actual results of a choice process are in hand that people may be able to pinpoint that they experienced a sense of manipulation. This suggests a connection between potentially regretting the outcome of a given choice and these two ethical concerns.

It is important to explore these ethical concerns in relation to combinations of digital nudges for both practical and theoretical reasons. On the practical side, charities need to know whether the combination of several nudges might have an adverse impact on potential online donors. Stacking several nudges intuitively entails more persuasive pressure on users [324]. However, backfiring effects have already been documented for single digital nudges (in the form of reminder emails) in the context of online fundraising [326]. There is also recent evidence that combinations of several nudges could potentially backfire [330], at least in the context of steering users towards more sustainable products. On the theoretical side, numerous empirical studies suggest a relevant relationship between users' sense of autonomy/manipulation and the efficacy of single pro-social nudges. To illustrate, Gråd et al. [321] compared defaults, social norms and moral commitment nudges to increase donations towards UNICEF. In this context, only some of the investigated nudges increased the proportion of users donating. Yet, the positive effects of the nudges (i.e. the increased tendency to donate) seemed to be driven by the users who did not perceive the nudges as attempts to manipulate their donation behavior, while donations among users who perceived the nudges as manipulative remained unaffected [321]. Additionally, their results indicated that the above nudges (which were compared to each other but not combined) did not crowd out charitable giving to any extent. In another study, default nudges seemed to undermine people's sense of autonomy when choosing between a hypothetical set of health insurance programs [335]. This applied to situations where there were only three health insurance programs to choose from, but when this set was increased to nine potential insurance programs, the drop related to people's autonomy was no longer significant. In contrast, Michaelsen et al. [336] conducted three experiments to examine threats to autonomy from default nudges. Their findings concluded that default nudges did not lead to a lower sense of autonomy or choice satisfaction on average. This remained true even when the presence of the default nudge was overtly disclosed to users [336]. So, while single nudges, especially defaults, seem not to be associated with threatened autonomy or a sense of manipulation, it is still unclear what happens when several nudges are combined in a pro-social context. This leads to our second exploratory research question:

RQ2: What is the relationship between the influence of pro-social digital nudge combinations and users' perceived threats to autonomy and sense of manipulation?

5.4 Digital Nudges for Charitable Giving

The related work suggests that digital nudges come in many different flavors and several literature reviews have characterized their design and use cases in research (cf. [77], [30], [73], [74]). We aim to combine their design properties systematically to evaluate their impact on charitable donations (RQ1) and then assess the relationship between these combinations and users' autonomy and sense of manipulation (RQ2). We will rely on three common patterns [73, 30]: default, friction, and social nudges and their subsequent combinations. The subsequent sections briefly outline how these nudges work, their design characteristics, and their application in previous studies.

5.4.1 Sticking with the status quo (default nudges)

Default nudges entail changing the choice outcome users will get by design if they do not explicitly request otherwise [103]. Several review articles across conventional and digital nudging have labeled defaults as one of the most effective tools for behavior change [337, 338, 73]. Psychologically, the reason defaults work is often explained by the status quo bias [339], meaning that when people are faced with choices, they tend to stick with the present state of affairs rather than change the outcome. This cognitive bias is, in turn, connected to loss aversion because potential disadvantages loom larger than the advantages of changing an outcome [340]. Therefore, it is preferable to stay with the status quo. In terms of design, default nudges shift the status quo in one or another direction and thereby leverage these psychological effects. In the digital sphere, defaults have been leveraged across many application domains from guiding drivers to find parking spots through default recommendations [341] to overcoming decision inertia related to investment choices in a study where users' earnings depended on the outcome of their investment decisions [106]. Also, several previous studies indicate that defaults alone can increase charitable giving contributions [342, 323, 343].

5.4.2 Toggling choice-related effort (friction nudges)

While changing the default option can strongly impact on choice outcomes, another element of design that can influence outcomes is to change the option-related effort - also known as friction [337]. As a nudge, adding or removing friction works because with all other factors held constant, humans tend to pick the path that involves the least effort or work [344]. Friction is a two-edged sword in terms of design because while it can guard against various user mistakes, it can also create user frustration and disengagement [345]. A classical example of how the application of friction can influence people can be found in a study on energy conservation behaviors. Houten et al. [346] encouraged people at a Canadian university to reduce the usage of elevators. They first tried various types of feedback and educational signs but these did not reduce the amount of energy consumed by the elevators. By instead introducing some friction to the environment (in this case an elevator door delay of about 11-16 seconds compared to the normal elevators) the researchers managed to reduce energy consumption by about 30 percent in all tested elevators as well as a reduction in the number of people using the elevators [346]. This would be an example of applying friction in terms of simply time delays. An example of friction being applied in the digital domain can be found in Kim et al. [65] who designed an app-level lockout task for smartphone users to discourage certain apps from being used. In a 3-week study, users were faced with the friction of inputting a random sequence of digits to access various apps in their phone related to Internet browsing, social media or entertainment. This type of design friction discouraged about half of users from accessing some of their phone apps [65].

5.4.3 Following the cues of the herd (social nudges)

In short, social design elements provide users with social reference points that reduce situational ambiguity and uncertainty [337]. While there are many nuances as to how these social norms are established through design, the tendency to follow the social cues of others seems to generalize across both online and offline decision contexts. Social nudges have been used to help users detect fake news stories on social media [347] to determine whether or not to accept a new kidney [348]. Previous studies suggest that social nudges can also interact with each other or other kinds of nudges in the sense that the efficacy of one of them can depend on the presence of the other (e.g. [105], [319], [86]). While social nudges have been successfully used in pro-social context, such as large-scale decreases in households' energy use [349], there are several examples where they fail as an intervention [350]. Previous studies suggest that descriptive vs. injunctive social norms tend to be more effective at changing behaviors because they imply less of a threat to autonomy and freedom than injunctive social norms [351]. Social norms are considered injunctive when they indicate what people should or should not do on moral grounds, and descriptive when they simply indicate typical behaviors of some relevant reference group and signal which behaviors are most popular [351].

5.5 Combining Defaults, Friction and Social Nudges

To investigate these nudge patterns and their combinations outside of the attitude–behavior gap we designed a simple online choice architecture for making an actual monetary donation to an actual charity, in our case, we picked the International Committee of the Red Cross (ICRC) [352]. In our study, donation choices involved real monetary donations for participants, sometimes referred to as a field task [343]. Figure 5.1 presents an instance of simple UI that contains no added nudges (control), and then a default, friction, and social nudge are added to the UI. For simplicity, we also decided to map potential donations as a percentage relative to the participants' earnings in the study. As Figure 5.2 will show, this creates a simple relationship between the amount donated and participants' earnings. We will now briefly describe the design and functionality of these digital nudges. The default nudge is simply a pre-selection of the 20% option indicated by its green color. If a user clicked *Confirm* when the default was present, the pre-selection will be their donation amount. The *Current Amount* was updated based on what the user selected, and when default nudges

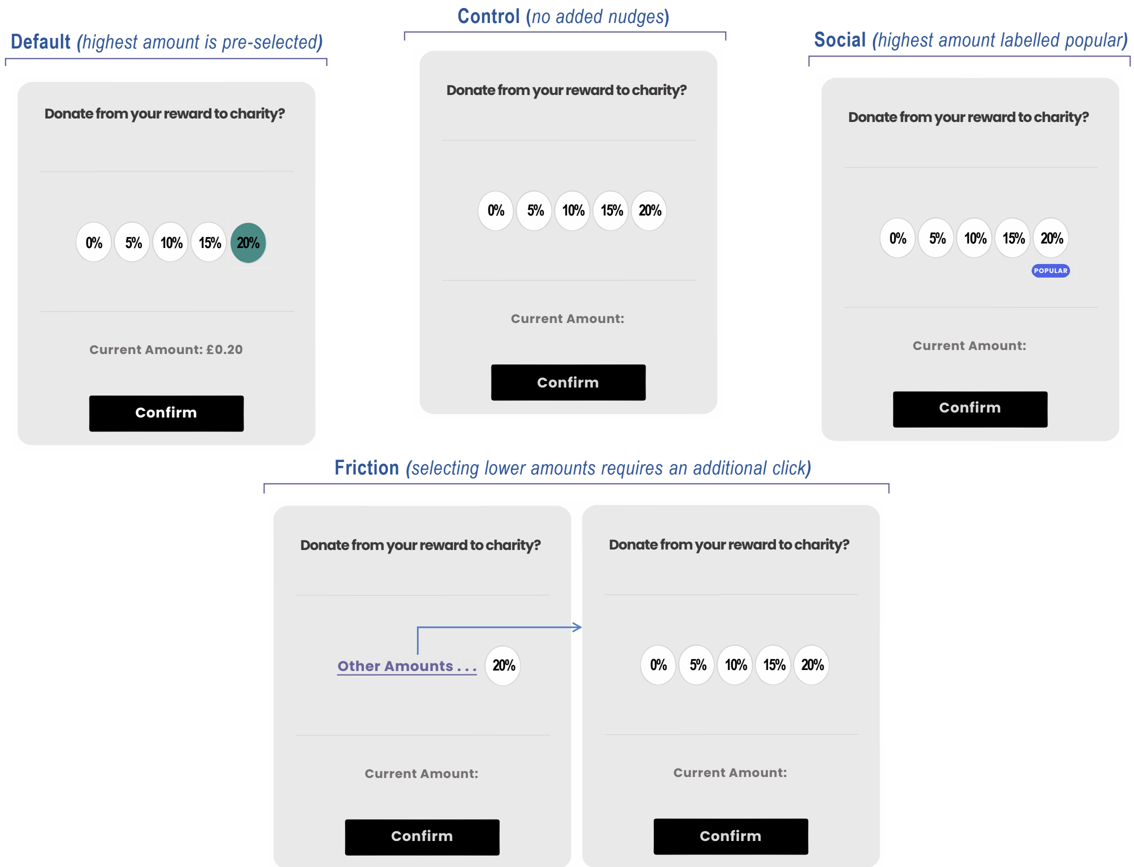

Figure 5.1: Mockups describing a donation UI with no added nudges (control) together with instances of default, friction and social nudges.

were present in the UI, this amount was pre-set to: £0.20 as in Figure 5.1. It should be noted that the user was only able to *Confirm* the donation with a single click if a default nudge was present. If no option was pre-selected or selected in the UI, and the user tried to click *Confirm*, the user would be prompted to again to select an amount. The friction nudge puts all the alternatives (except the 20% option) two more clicks away. In other words, if friction is present and the user wants to donate less than 20%, they will need to first click *Other Amounts...*, and then select their desired donation option as shown in Figure 5.1. Finally, the social nudge is the blue badge under the 20% option, labeling it as the *Popular* choice. Note that all the digital nudges were pointed in the pro-social direction (i.e. to maximize the donation amount to the charity) *. That is, users are steered by design towards donating 20% of their rewards.

5.6 Evaluation

To explore our two research questions, the evaluation was designed as an online factorial experiment where participants were randomly allocated to one of 2^3 possible conditions (default: on vs off; friction: on vs off; social: on vs off) thus involving 0, 1, 2 or 3 digital nudges combined. Figure 5.2 shows all of the eight randomized nudge patterns that were used in the evaluation. The UIs were made with JS and CSS modifications directly in Qualtrics. Before the main data collection, we also conducted three pilot benchmarks ($n = 104$, $n = 102$, $n = 101$) with the UI in Figure 5.2 A), where we tested different amounts as donation options. This was to ensure that the donation amounts were properly scaled as to avoid a context where everyone would give away all of the pre-set portion, or a situation where nobody would give away anything. Based on these pilots, we settled on a design where the amounts ranged between 0% – 20% because the average and median donation amount in this design was close to the middle option. If this is used as a benchmark, it allows for both increases and decreases in donation amounts.

*We realize that this may stretch the definition of what traditionally is labeled as a nudge. Hagman et al. [318] outline that pro-social nudges aim to improve overall social welfare, even when doing so is inconsistent with maximizing private welfare.

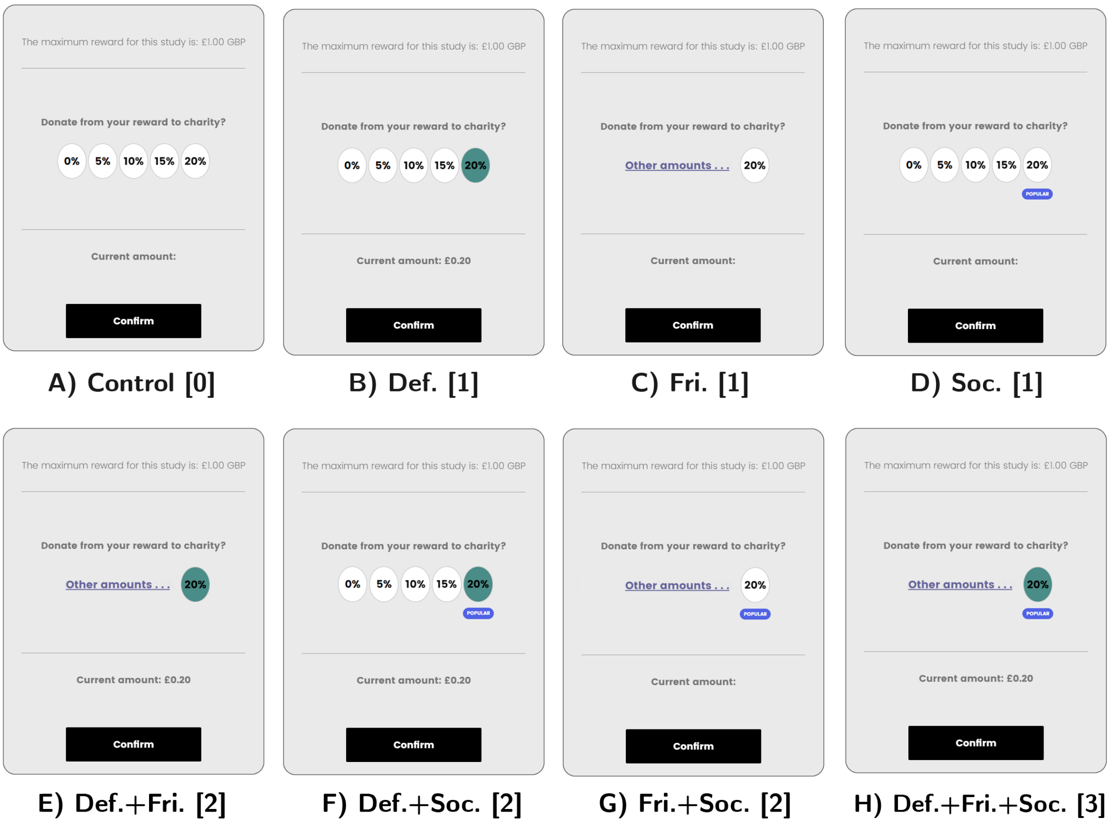

Figure 5.2: The eight nudge patterns used for the experimental groups combining default (Def.), friction (Fri.) and social (Soc.) nudges. The number of nudges present in each UI are in brackets.

5.6.1 Participants

Our aim was to gather one hundred participants for each of the eight experimental groups. The participants for the experiment were drawn from Prolific. Participants from the pilot benchmarks could not participate in the main study. From Prolific, the participants were directed to an anonymous Qualtrics link. The participants were instructed that in the first part of the study, they could donate part of their reward to the ICRC, and that the second part was focused on the design of charity apps. At this stage, the participants were not aware that the study was focused on digital nudging. The Prolific participants were selected using the following screening criteria:

UK sample, Approval Rate >= 95%, usage of a laptop/tablet and Google Chrome web browser. The last screening criteria were used to minimize device and browser compatibility issues related to the UI designs. Each study participant was compensated £0.80-1.00 GBP for their participation (depending on their donation to the ICRC). While modest, the pilot tests and the main study raised a total of 96.44 GBP, subsequently donated to the ICRC.

5.6.2 Procedure & interaction

After the informed consent screen, an attention check was deployed, and participants were given a commitment request based on the example from Geisen [353]. On the subsequent screen, participants received a short introduction to the work of the ICRC as a charity. The exact information that participants saw on this screen can be found in the Appendices section.

After this information, participants were randomized to one of the UIs in Figure 5.2, where they could select a donation amount. After the interaction with one of the UIs in Figure 5.2, participants were asked to respond to a series of scale items. First, participants were asked about perceived threats to their autonomy [354], and perceived sense of manipulation [355]. Participants were explicitly told that these two scales related to the design of the options that they were given when choosing the donation amount on the previous screen. At this stage, participants were also asked to rate the UI design that they had been exposed to through the Standardized User Experience Percentile Rank Questionnaire (SUPR-Q)[356] *.

*Unfortunately, due to a data capture error, we failed to record one of the scale items related this measure, we therefore cannot compute a valid SUPR-Q score.

Participants were then asked about their attitudes towards charitable organizations [357] before finally answering demographic questions. In this last section of the study, participants were reminded of the amount they had decided to donate, and faced with the choice of either sticking with or changing their donation amount.

5.6.3 Measures

We will now describe the measures used for the evaluation in more detail.

Donations

This measure was operationalized on two different levels to capture various aspects of users' donation behaviors.

Donation amount. This pertains to the button option a given user pressed (0%–20%) and then confirmed in the UIs, as seen in Figure 5.2 and 5.1.

Compliance. Operationalized as a dichotomous variable (0/1), one means that the user was steered in the direction of the nudges and donated the highest amount i.e. 20%. A zero means that the user took another path then what the digital nudges suggested (0%–15%). From this, a compliance percentage can be derived for each of the nudge patterns.

Attitude towards charitable organizations (ACO)

The items for this scale were adapted from Webb et al. [357]. These five items constitute one of two dimensions of a larger construct aimed at determining people's propensity to donate to charitable organizations[358]. In our study, this variable is used as an attitudinal control measure. In a random presentation order, users responded to these items on a 7-point Likert-scale ranging from "Strongly Disagree"

(1) to "Strongly agree" (7). An example item from this scale would be: "My image of charitable organizations is positive." The internal consistency reliability of this scale was (Cronbach's $\alpha = .908.$).

Perceived threat to autonomy

The items for this scale were adopted from Dillard and Shen [354]. Originally, this scale was referred to as "Threat to freedom" [354], but in line with other papers on persuasive design (cf. [359]) we have opted for the current label. Users where presented with a 7-point Likert-scale on each item ranging from "Strongly Disagree" (1) to "Strongly agree" (7). We revised our items focusing on the potential threat from the "design" rather than the "message" [354]. To illustrate, an item from this scale could be: "The design threatened my freedom to choose." The presentation order for these items was randomized. In our study, the internal consistency reliability of this scale was acceptable (Cronbach's $\alpha = .878$).

Perceived sense of manipulation

The items for this scale come from Witte [355]. Users were presented with a 7-point Likert-scale, where we again revised our items to focus on the design rather than a message. A concrete item from this scale would be: "The design tried to manipulate my feelings". The presentation order for these items was also randomized. The Cronbach's α coefficient for this scale was .915.

Changed donation decision

Towards the end of the study, all users were reminded of their previous donation amount, and explicitly asked if they wanted to change their previous decision to donate this particular amount. This measure was also dichotomous (0/1) where zero meant that the user stuck with their previous decision to donate a certain amount, and a one means that the user changed their previous donation amount. We will use this action as a proxy measure for regret, Eyal [360, p. 1] refers to this as a "regret test" for users.

User experience

Finally, users who were exposed to one of the UIs containing digital nudges (Figure 5.2 B– 5.2 H) were asked to textually describe the difference between the UI they used for their donation and the UI containing no digital nudges (Figure 5.2 A). Images of the two Uis were shown in random presentation order to users.

5.6.4 Evaluation results

The following section contains three main parts. First, we provide descriptive statistics related to our evaluation. Second, we explore our first research question on whether the pro-social nudge combinations have cumulative or interactive effects on users' charity donation behaviors. The third section examines the second research question regarding the relationship between the influence of the pro-social nudge combinations and users' perceived threats to autonomy and sense of manipulation. We also provide a qualitative content analysis of users' textual descriptions of the nudge combinations. Due to the distribution of some of our measures, we will later use non-parametric tests for most of our inferential conclusions *.

*The reporting of the inferential results will conform to the statistical recommendations of significance described by Benjamin et al. [234].

Descriptive statistics

In total, 805 participants completed the main study. Out of these, we removed three participants who failed the attention check, and eight participants because their browsers were not linked to a desktop device (e.g. iPhones), resulting in 794 valid responses. It took participants approximately 5 minutes to complete the study ($\hat{m}$ = 4.95, $\hat{\sigma}$ = 3.13). The sample consisted of 49.9% females, and participant ages ranged between 18 and 78 years old ($\hat{\mu}$ = 41.42, $\hat{\sigma}$ = 12.82). The randomization procedure was deemed successful because we detected no systematic differences regarding sex, age, or ACO scores between the experimental groups that would not be expected by chance. In the next sections, we will control for the latter two variables in our inferential analyzes but first we will descriptively explore the donation amounts across all the nudge patterns. Figure 8.2 in the Appendices section shows the distributions and associations of the continuous measures in the study. Table 5.1 shows the nudge pattern, number of nudges present in the UIs, the proportion of users who complied with the given nudge pattern, and four summary statistics for the donation amount.

Table 5.1: Descriptive statistics related to compliance percentage and donation amounts across all the nudge patterns.

			Compliance	Donation Amount			
Nudge patterns	No. of nudges	n	(%)	$\sum$	$\hat{m}$	$\hat{\mu}$	$\hat{\sigma}$
Control (no nudges)	0	95	28.42	830	10.00	8.74	8.19
Default (Def.)	1	97	32.99	870	5.00	8.97	8.57
Friction (Fri.)	1	98	50.02	1205	20.00	12.30	8.56
Social (Soc.)	1	102	22.55	790	5.00	7.75	8.10
Def. + Fri.	2	100	54.00	1195	20.00	11.95	9.24
Def. + Soc.	2	102	36.27	970	10.00	9.51	8.80
Fri. + Soc.	2	100	39.00	975	10.00	9.75	8.92
Def. + Fri. + Soc.	3	100	63.00	1340	20.00	13.40	8.96

Summary statistics for sums ($\sum$), medians ($\hat{m}$), averages ($\hat{\mu}$) and standard deviations ($\hat{\sigma}$)
All numbers have been rounded to two decimals, and the unit of the sum is British pence (£/100)

(RQ1) The impact of nudge combinations on donation behavior

The average donation amounts are visualized when the data are aggregated across the number of nudges in the UIs (i.e. 0, 1, 2, 3) in Figure 5.3 A). This figure suggests an increase in donation amounts when the number of nudges increases in the UI, especially when comparing three nudges to the control group. Figure 5.3 B) shows the average donation amounts across each specific nudge pattern in the UIs (i.e. default, friction, and social). The main trend in Figure 5.3 B) is that donation amounts seem to increase when some combinations of friction are present in the UI. The whiskers in all subsequent barcharts indicate 95% mean confidence intervals.

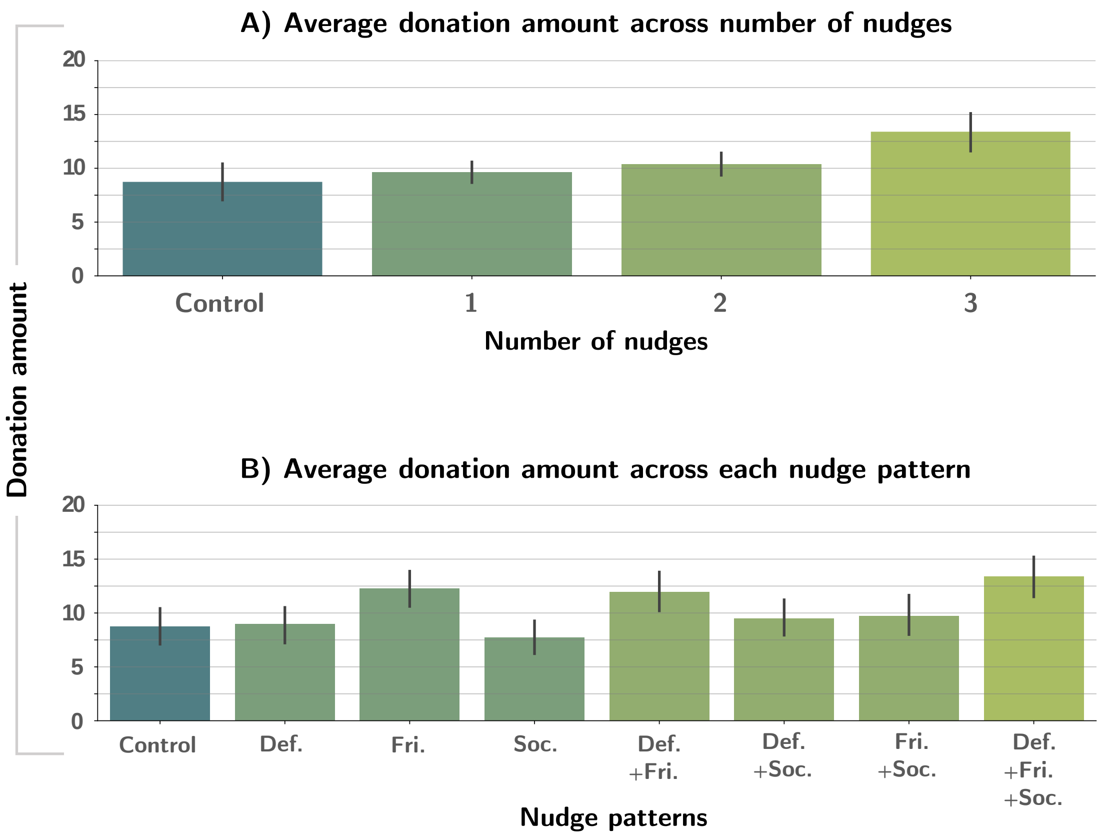

Figure 5.3: Barchart A) showing the average donation amount across the number of nudges stacked in the UI. B) showing the average donation amounts across each specific nudge pattern ($n = 794$).

Likelihood to comply with the nudge combinations To further test the patterns shown in Figure 5.3, we specified three multiple logistic regression models with compliance (0/1) as the dependent variable*. All the models control for users' ages and ACO scores. These two predictors have been mean centered to aid interpretation. In the first model, we added the number of nudges as dummy variables with indicator coding. This model connects to the results shown in Figure 5.3 A). Model 2a and 2b connect to the results shown in Figure 5.3 B). In Model 2a, the number of digital nudges was exchanged for every specific nudge pattern. Model 2b complements Model 2a by accounting for the subsequent interaction terms of the nudge combinations. Note that the adjusted odds ratios ($OR_{Adj.}$) for the dummy variables in Model 1 and 2a are estimated relative to the control group, whereas the $OR_{Adj.}$ for the dummy variables in Model 2b are estimated relative to the presence of each nudge combination, thereby accounting for the factorial experimental design. Table 5.2 provides information about the contribution towards compliance for each predictor variable in the models. The model diagnostics results are shown in Table 8.2 in the Appendices section. The VIF-values between the predictors in all models were deemed unproblematic based on established recommendations for multicollinearity [361, 362].

The first model indicates that, on average, when only one nudge is added (regardless of what it is), the odds of complying (donating the highest amount) relative to the control group do not change. Two nudges have a suggestive effect, with users being almost twice as likely to comply with the nudges ($OR_{Adj.} = 1.944$, $W = 6.242$, $p = .012$) compared to the control group. When three nudges are added, the odds of complying compared to the control group is almost four times higher ($OR_{Adj.} = 3.898$, $W = 18.128$, $p < .001$).

*It can be argued that the most relevant measure in a donation context would be the donation revenues i.e., the sums. The correlation between the average compliance rate and the donation sums in Table 5.1 is (r = 0.989) and we, therefore, view compliance as a good proxy for donation revenues.

Table 5.2: Predictors for the likelihood to comply with the digital nudges (i.e. donate the highest amount) in the multiple logistic regression models.

	Regression model parameters			
	$\widehat{\beta}$	‡$OR_{Adj.}$	[95%CI]	Wald
Model 1: Reference category for each dummy variable is the control group				
Control (β_0)	-0.985^{***}	0.397		17.269
Age	0.024^{***}	1.024	[1.012, 1.036]	15.502
ACO score	0.488^{***}	1.629	[1.418, 1.872]	47.261
†Number of nudges : 1	0.373	1.452	[0.858, 2.458]	1.934
†Number of nudges : 2	0.665^{*}	1.944	[1.154, 3.274]	6.242
†Number of nudges : 3	1.360^{***}	3.898	[2.084, 7.292]	18.128
Model 2a: Reference category for each dummy variable is the control group				
Control (β_0)	-0.992^{***}	0.371		17.379
Age	0.023^{***}	1.024	[1.011, 1.036]	14.377
ACO score	0.523^{***}	1.687	[1.460, 1.950]	50.130
†Default	0.282	1.326	[0.697, 2.522]	0.740
†Friction	1.026^{**}	2.791	[1.496, 5.205]	10.416
†Social	−0.290	0.749	[0.382, 1.466]	0.713
†Default + Friction	1.225^{***}	3.406	[1.826, 6.353]	14.842
†Default + Social	0.290	1.336	[0.711, 2.509]	0.810
†Friction + Social	0.460	1.584	[0.845, 2.970]	2.058
†Default + Friction + Social	1.358^{***}	3.889	[2.075, 7.291]	17.945
Model 2b: Reference category for each dummy variable is its own absence				
Control (β_0)	-0.992^{***}	0.371		17.379
Age	0.023^{***}	1.240	[0.671, 2.293]	50.130
ACO score	0.523^{***}	1.240	[0.671, 2.293]	47.261
†Default	0.282	1.326	[0.697, 2.522]	0.740
†Friction	1.026^{**}	2.791	[1.496, 5.205]	10.416
†Social	−0.290	0.749	[0.382, 1.466]	0.713
†Default × Friction	−0.083	0.920	[0.387, 2.191]	0.035
†Default × Social	0.297	1.346	[0.540, 3.351]	0.407
†Friction × Social	−0.277	0.758	[0.310, 1.857]	0.367
†Default × Friction × Social	0.402	1.495	[0.434, 5.144]	0.407

* $p < 0.05$, ** $p < 0.005$, *** $p < 0.001$
†Dummy variable with indicator coding (e.g. 0/1)
‡Adjusted odds ratios, with all other variables held constant in the model

The second model corroborates the results seen in Figure 5.3 B). The model highlights that users who were exposed to three of the nudge patterns were significantly more likely to comply with these nudge combinations and thereby donate higher amounts on average, see Table 5.2. This was true for Fri. (shown in Figure 5.2 C), Fri.+Def. (Figure 5.2 E), and Fri.+Def.+Soc. (Figure 5.2 H).

However, the common denominator suggested by both Figure 5.3 B) and Model 2b is that friction alone is the driving factor of increased donations. Model 2b shows that this is the only significant dummy predictor, indicating that users who were exposed to a nudge pattern that contained friction were almost three times more likely to comply ($OR_{Adj.} = 2.791$, $W = 10.416$, $p = .001$), while controlling for other factors. Furthermore, Model 2b suggests no interactive effects from any particular combination of the default, friction and social nudges related to compliance.

(RQ2) Relationship between nudge combinations, autonomy and manipulation

Figure 5.4 A) shows the relationship between users' perceived threats to autonomy collapsed across the number of nudges. The previous section showed that compliance percentage and donation amounts go up, especially when two or three nudges are present. Figure 5.4 highlights that the potential increase in donation amounts comes at the price of perceived threats to autonomy. This is born out on two levels. First, Figure 5.4 A) shows that as soon as nudges are present in the UI, the threat to autonomy goes up on average. A Mann-Whitney U-test indicates a significant difference in perceived threat to autonomy between the control group (i.e. no nudges) ($\hat{m} = 2.25$, $n = 95$) and the pooling of all the groups that were exposed to various nudges ($\hat{m} = 4.25$, $n = 699$), $U = 17349.0$, $z = -7.558$, $p < 0.0005$, $r = 0.477$.

Second, Figure 5.4 B) shows the relationship between compliance and the perceived threat to autonomy but now collapsed across all of the nudge patterns. The overall trend here is that the threat to autonomy is especially high among users who

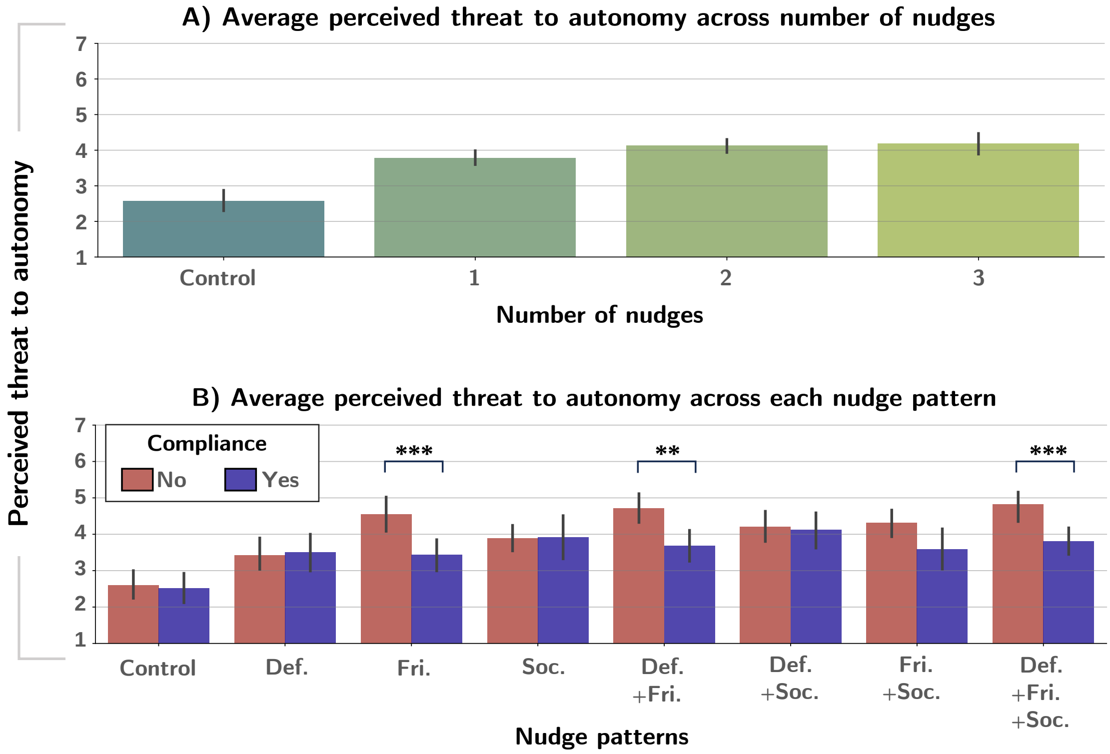

Figure 5.4: Barchart A) shows the average threat to autonomy across the number of nudges and B) shows the average threat to autonomy across each nudge pattern clustered on compliance ($n = 794$). Mann-Whitney U-tests indicate differences between the compliance clusters (No/Yes) within each nudge pattern (* $p < 0.05$, ** $p < 0.005$, *** $p < 0.001$). These tests were performed without using a family-wise error correction, i.e., $(\alpha_E/1)$.

do not comply with some nudge patterns. As a follow-up, we looked at the differences in perceived threats to autonomy for users who complied vs. those who did not comply with each of the nudge patterns, and the results of the Mann-Whitney U-tests where we detect differences are shown in the figure.

Figure 5.4 B) suggests a higher perceived threat to autonomy among users that did not comply with the following nudge patterns: Fri., Fri.+Def., and Fri.+Def.+Soc.. The Appendices section contains additional box plots that visualize the results for all eight Mann-Whitney U-tests.

Users' sense of manipulation is tightly associated with their perceived threats to autonomy. Figure 5.5 A) highlights the relationship between users' sense of manipulation across the number of nudges. Figure 5.5 A) shows the same trend as before: As soon as nudges are present in the UI, the sense of manipulation increases on average. Another Mann-Whitney U-test also indicates a significant difference for sense of manipulation between the control group ($\hat{m} = 2.00$, $n = 95$) and all the groups that were exposed to various nudges ($\hat{m} = 3.33$, $n = 699$), $U = 22952.0$, $z = -4.887$, $p < 0.001$, $r = 0.309$. Figure 5.5 B) shows the relationship between compliance and the sense of manipulation collapsed across all of the nudge patterns. As a follow-up, we again looked at the differences in the sense of manipulation for users that complied vs. did not comply with each of the nudge patterns, and the results of these Mann-Whitney U-tests are shown in Figure 5.5 B).

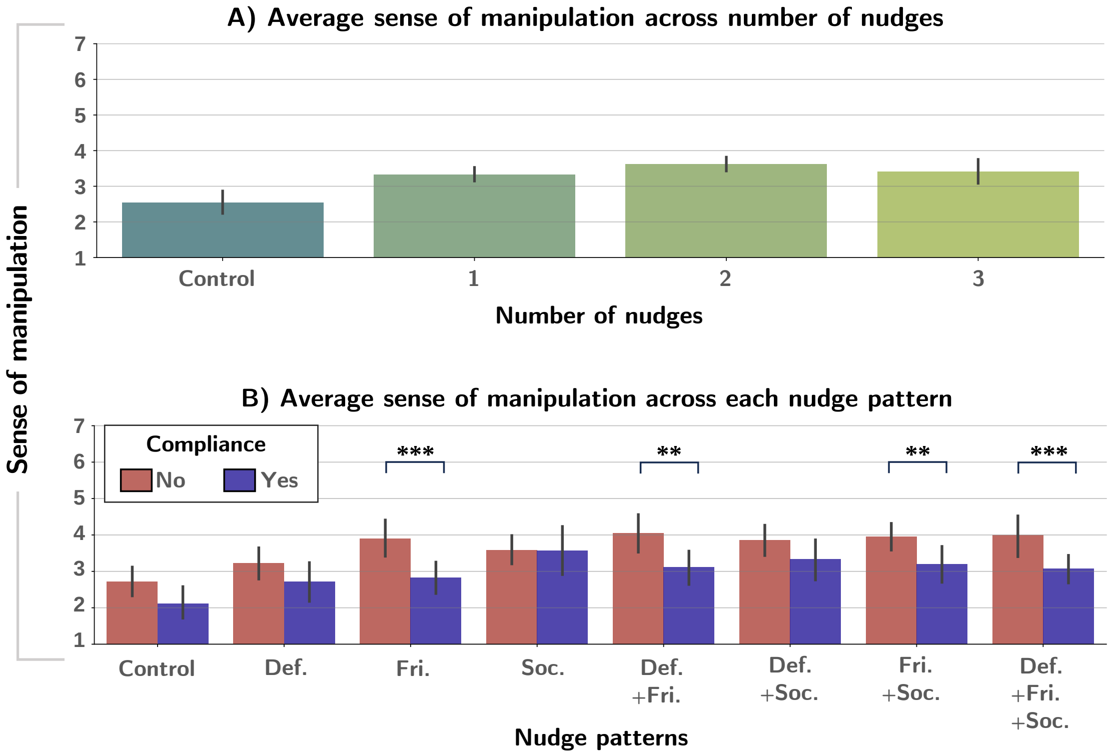

Figure 5.5: Barchart A) shows the average sense of manipulation across the number of nudges and B) shows the average sense of manipulation across each nudge pattern clustered on compliance ($n = 794$). Mann-Whitney U-tests indicate differences between the compliance clusters (No/Yes) within each nudge pattern (* $p < 0.05$, ** $p < 0.005$, *** $p < 0.001$). These tests were performed without using a family-wise error correction, i.e., $(\alpha_E/1)$.

A similar trend emerges – the sense of manipulation is particularly high among users that do not comply with different combinations of friction nudges. The Appendices section contains additional box plots that visualize the results for all eight Mann-Whitney U-test results related to sense of manipulation.

Finally, across the whole sample, only 27 users ($\approx 3.4\%$) changed their donation amounts at the end of the study. There was no suggestive association between the nudge patterns and this decision χ^2 $(df = 7, n = 794) = 1.827, p = 0.969$. While the number of users here is admittedly small, when the donation amount was changed, there are clear differences between users who changed their donation and those that did not in terms of perceived threat to autonomy and sense of manipulation, see Figure 5.6 below.

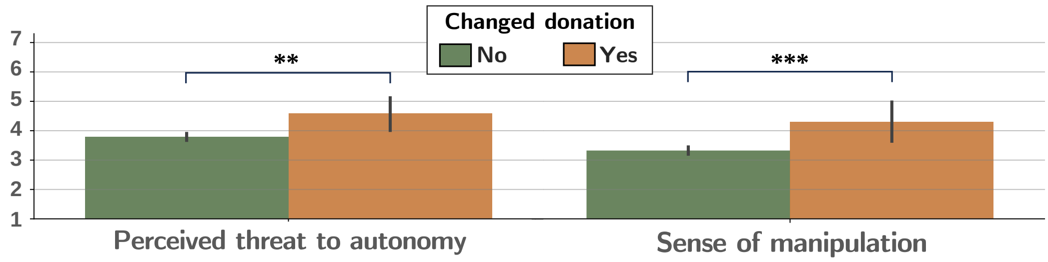

Figure 5.6: Barchart shows the average perceived threat to autonomy and sense of manipulation clustered on those who stayed with their donation decision (No, $n = 767$), and those who changed their donation amounts (Yes, $n = 27$). Mann-Whitney U-tests indicate differences between these clusters (No/Yes) within each scale (* $p < 0.05$, ** $p < 0.005$, *** $p < 0.001$).

Qualitative User Comments Analysis

In this section, we take a more detailed look at the user experience associated with the nudge patterns. Our aim is to get a better understanding of users self-reported perceptions of the different nudge patterns. It should be noted that only users who were exposed to the digital nudges (Figure 5.2 B– H) were asked to describe the difference between the UI they interacted with and the UI containing no nudges (Figure 5.2 A).

Out of those users, everyone provided answers ($n = 699$). We will now present an exploratory content analysis of the responses to this open-ended question. This content analysis was conducted through a systematic process involving two of the researchers. First, both researchers independently read through all comments once, and manually coded 70 randomly selected comments ($\approx 10\%$), then engaged in discussions to identify initial codes [363]. The coding process was conducted via a shared spreadsheet. A first set of 39 agreed codes was then independently applied to another 70 randomly selected comments by each researcher. The inter-rater reliability between the two researchers was assessed after this initial round using Cohen's Kappa ($k = 0.648$), indicating moderate agreement [364]. After the initial coding, four of the co-authors gathered to (1) combine similar codes into more coherent ones and (2) identify additional codes to better capture users' descriptions. For example, we merged the codes `pushy` and `forced` into a single code `pressure` and introduced the `nudge_preference` code. After this process, our codebook included a total of 21 codes. The two researchers who initially coded the comments now coded all comments together using the refined codebook. At the end of the process, the two researchers conducted a final round of independent coding using the final codebook by re-sampling another 70 random comments. Cohen's Kappa was again calculated (now at $k = 0.803$), signifying almost perfect agreement. Finally, four of the researchers engaged in discussions to establish the main themes. These themes are described in the subsections below together with associated comments from users.

Feeling exploited by the design. The first theme that emerged from our analysis relates to a feeling of trickery and manipulation expressed by many participants, most of whom were exposed to nudge combinations with friction. Some participants simply stated they felt manipulated by the nudge. For instance, participant P87 (Male, 47)$_{\text{Def.+Fri.}}$ wrote: *"the [Control] one leaves you to make your own decision. the [Def.+Fri.] one tries to manipulate you."*, and P421 Female (53)$_{\text{Fri.+Soc.}}$ stated: *"[Fri.+Soc.] is manipulative and attempting to force donations."*. Others associated the feeling of manipulation with a reduction of choice caused by the nudge. For example, P614 Female (33)$_{\text{Fri.}}$ noted: *"Design [Fri.] feels a bit more manipulative in that it makes you feel like 20% is the amount most other people choose by making it the only option viewable on the screen. [...]."* and P623 Female (49)$_{\text{Fri.}}$ wrote: *"[Fri.] only displaying the largest % that can be donated this will impact some users who are not able to understand if they have options"*. Finally, some participants also pointed out the potential backfire effect of employing such a technique in a charitable giving context. For instance, participant P127 Male (52)$_{\text{Fri.}}$ stated: *"The [Fri.] design is counting on people not working out how to change the amount donated - I find it covertly manipulative. I donate a lot of money to charity but through my own free will rather than as a reaction to pressure."* and P627 Female (58)$_{\text{Def.+Soc.}}$ commented: *"The popular feature is often used by online stores such as VPN/antivirus sellers to steer a user into buying a more expensive product. I'm not sure if a charity would want to resort to that unless really desperate for funds. It might raise a bit more but at the expense of reputation."*.

Reactions to perceived pressure. The second theme that emerged from the analysis reflects participants' reactions to the pressure they felt from the nudges. It is worth noting that most of these participants were in the Social group. Some participants stated that they felt pressured by the nudge, with for example participant P578 Female (63)$_{\text{Soc.}}$ commenting: *"Minimal, but pressurising"* and P196 Male (21)$_{\text{Soc.}}$ writing: *"I think is a bit too pushy for my liking."*. Some participants linked the pressure to a sentiment of guilt for not being able to comply with the nudge. For example, P695 Male (47)$_{\text{Soc.}}$ wrote: *"I think [Soc.] could make people feel guilty if they weren't able to commit to that amount, the feeling of guilt could make people uncomfortable enough to leave with-*

out pledging anything.”, and P243 Male $(36)_{\text{Fri.}}$ stated: *“[Control] lets you pick how much without any pressure at all, where as [Fri.] selects the highest amount automatically you feel a bit guilty changing it lower.”*. P624 Female (52) $_{\text{Soc.}}$ pointed out the fact that the nudge allows for less freedom of will: *“[Control] lacks the ”popular” tag, allowing more discretion and expression of free will than [Soc.].”* Finally, some other participants questioned the truthfulness of having the highest amount selected, with for example P72 Female $(35)_{\text{Def.+Soc.}}$ commenting: *“[Def.+Soc.] leaves you puzzled if its true that 20 percent is most popular donation. You feel pressured to donate.”* and P450 Male $(41)_{\text{Soc.}}$ noting: *“suggesting a random number is popular is not credible. Best to leave that out and allow the user to decide what they can afford”*.

Dispassionate description of the design. Another theme that was generated from analyzing participants’ comments comprises dispassionate descriptions where they simply described the visual appearance of the interfaces. It is worth noting that most of these participants were in the Default group. In this regard, participant P375 Male $(75)_{\text{Def.}}$ indicated: *“[Def.] is suggesting the maximum”*, and participant P456 Female $(27)_{\text{Def.}}$ wrote: *“One has the 20% option pre-selected, while the other doesn’t have any option selected”*. A small proportion of participants from the Social group simply described the popular option under the nudge, with for example participant P286 Female $(45)_{\text{Soc.}}$ mentioning *“Design [Soc.] has an option to see popular donations”* In the Friction group, participants described the nudge pattern as limiting their choice, with for example participant P287 Female $(30)_{\text{Fri.}}$ noting: *“The [Control] design gives you the options straight away without having to look for them”*, and participant P531 Male $(49)_{\text{Fri.}}$ writing: *“For [Fri.] you have to physically click on the donation amount to change it, but the other has the different amounts to just choose.”*. When several nudges are combined, participants usually do not describe all of them; instead, they tend to focus on a specific nudge.

For example, participant P95 Female $(43)_{\text{Def.+Fri.+Soc.}}$ wrote: *"[Control] gives you options straight up the other you need to click on 'Other Amounts' to see what other options are available"*, and participant P416 Male $(27)_{\text{Def.+Soc.}}$ commented: *"[Def.+Soc.] indicates the most popular reward to charity advising to pick the highest percentage"* and participant P274 Female $(41)_{\text{Def.+Fri.}}$ noting: *"The donation portion is different, [Control] shows all the percentages the other doesn't"*.

Choice facilitation. Several participants, although a minority, mentioned that the nudge they were exposed to facilitated their donation choice. For instance, participant P137 Female $(44)_{\text{Def.+Fri.+Soc.}}$ expressed finding the nudge pattern simple and helpful: *"The [Def.+Fri.+Soc.] option offers a suggestion on how much to donate, as a user I find this to be helpful rather than manipulative. [Def.+Fri.+Soc.] options have simple and easy to use design."*, and participant P153 Male $(38)_{\text{Def.+Fri.+Soc.}}$ wrote: *"The [Def.+Fri.+Soc.] is simple and direct, using it is easy. Design [Control] has more options on the screen so you feel you have more choice.*. Other participants stated that they preferred the nudge design because it offered a convenient amount to donate, such as participants P124 Female $(53)_{\text{Def.+Fri.+Soc.}}$ who wrote: *"[The control] provides user with greater level of freedom to make a decision. I prefer [Def.+Fri.+Soc.] as the default position is an appropriate level of donation to make"*. Another participant, P140 Female $(35)_{\text{Def.+Fri.+Soc.}}$, stated they preferred the nudge design because they found it more decisive: *"i think option [Def.+Fri.+Soc.] feels more convenient, its more decisive. I would imagine the longer you have time to make your choice perhaps the less you would donate."*. Additionally, participant P221 Female $(29)_{\text{Def.+Fri.+Soc.}}$ explained the found the nudge to be more appealing: *"the [Def.+Fri.+Soc.] one is user friendly as it has real money amount in the percentage so that you dont have to sit and work it all out. [...]. the [Def.+Fri.+Soc.] one is way more appealing"*.

The themes regarding users' self-reported perceptions of the nudge combinations will be discussed together with the other findings of the evaluation in the next sections.

5.7 Discussion and Implications

This chapter explored the additive and interactive effects of digital nudge combinations in a charitable giving context. Apart from just measuring the behavioral impact from the nudge combinations, we echoed the point that it's important to consider additional perspectives from users concerning persuasive technologies [336, 327, 328]. We also explored the relationship between these pro-social nudge combinations and users' perceived threats to autonomy and their sense of manipulation. Additionally, our qualitative analysis of user comments unveiled valuable insights about their perceptions of these nudge patterns.

5.7.1 Nudge combinations and donation behavior

When aggregating users' donations on the number of nudges present in the UI, our results suggest that stacking 2-3 nudges does have an additive impact on donations and users' tendency to comply with the nudges when compared to no added nudges (control). More specifically, in our case, three nudge patterns Fri., Fri.+Def., and Fri.+Def.+Soc. significantly increased the likelihood to comply which increased the subsequent donation amounts compared to the control, see Table 5.2. Yet, once we account for interaction effects, our results suggest that the friction nudge (putting the lower donation options only two clicks away) is the driving factor of increased donations. This point is illustrated in Figure 5.3, and Model 2b in Table 5.2.

In the related work, we mentioned that previous research has often investigated one nudge at a time in the context of charitable giving. Our study advances knowledge in this area by accounting for the potential interactions between three pro-social nudges. Our findings also suggest that two of our proposed nudges (default and social) did not impact donations on their own.

Concerning default nudges, previous research shows mixed results in similar contexts. On the one hand, Altmann et al. [320] found that aggregate donation levels concerning rather small amounts of money (where the average donation was less than 2 euros) were unaffected by defaults in a large-scale field experiment. Goswami and Urminsky [365] also found in one of their field studies that there were no effects on average donation amounts of setting the high or medium donation option as the default. On the other hand, Ghesla et al. [342] found that higher default amounts led to significantly higher donations. Another study found that 81% of users donated half of their earnings to charity when this was set as the default option in the experiment (compared to 19% when keeping their earning was the default) [323]. Another study found that opt-out defaults increased donations by about 25 percent compared to opt-in defaults [343]. In our study, the default nudge only had an impact relative to the control group when it was combined with friction, but our findings suggest that the friction nudge is what drives this impact.

Our social nudge even shows some signs of a backfiring effect. While we cannot detect significant differences between this nudge (alone) and the control group, we observe that friction increased donations on its own but when it was combined with the social nudge this effect was diminished. A recent review on backfiring effects from behavioral interventions suggested that social nudges were the most common intervention resulting in failures [350]. Our qualitative comment analysis sheds further light on why this might be. Indeed, some participants expressed signs of reactance against the direction in which the nudge was attempting to steer them once they realized they were being nudged. Others questioned the truthfulness of the social nudge, highlighting a lack of credibility in what was shown to be the popular donation option. While their application domains are different, there are at least two recent studies that detected interaction effects between default and social nudges (see [105], [115]). Shore and Cummings [319] did not detect such an interaction in the context of online privacy, and neither did we. Another recent study even suggests a backfire effect from these two nudges when users were steered towards green fashion products [330]. This further suggests that more research needs to be conducted where digital nudges are systematically combined.

The current study demonstrates the possibility for online charities to stack several pro-social nudges. Previous researchers have argued that employing nudges is ethically permissible, even when the autonomy of the potential donor might be threatened, if this alleviates the suffering of victims affected by life-threatening events [317] [*]. While many charities can faithfully employ such an argument, our findings suggest that great care should be taken when designing combinations of pro-social nudges. Friction in particular seems to impact charitable donations, but the advantage of these nudge combinations should be understood in light of their side effects pertaining to perceived threats to autonomy, sense of manipulation and users' self-reported comments.

5.7.2 Nudge combinations, autonomy and manipulation

Our second research question concerned the relationship between the influence of these nudge combinations and users' perceived threats to autonomy and sense of manipulation. Previous researchers have argued for the importance of measuring direct user experiences related to autonomy and manipulation when nudges are em-

[*] This argument certainly applies to charities that help victims of armed violence or war, such as the ICRC.

ployed[336]. To answer this call, we tried to incorporate users' perspectives on these two ethical concerns in our evaluation. Two main points from our findings seem to emerge. First, users detect these pro-social nudges both in terms of their own perceived autonomy and sense of manipulation. On average, introducing these pro-social nudges comes with a heightened threat to autonomy, and sense of manipulation.

The least problematic nudge pattern in terms of autonomy or manipulation from users perspectives seems to be the default nudge (alone), which is in line with some previous research [336, 321]. Nudge combinations that involve social or friction elements seem more problematic though. Our qualitative analysis offers perspectives for understanding the increased sense of manipulation resulting from these nudges. While some participants stated that they felt manipulated because the nudges reduced their freedom of choice, others felt pressured due to feelings of guilt or shame induced by the nudges. This offers important implications for online charities that employ pro-social nudges, because they may backfire in certain cases, resulting reactant behaviors [350, 366].

The second point is that the nudge patterns that work in terms of changing donation behaviors seem to have polarizing effects on users' perceived threats to autonomy and sense of manipulation. Users who were exposed to the three nudge patterns (Fri., Fri.+Def., and Fri.+Def.+Soc.) consistently reported a higher perceived threat to their autonomy when they resisted these nudges. For users' sense of manipulation, we detect a similar signal but this signal is not as clear as for the threat to autonomy. As the Appendices section shows, when we correct for multiple comparisons, we no longer detect differences for the sense of manipulation between users that complied and those that did not in the Fri.+Def., and Fri.+Def.+Soc.) groups. Yet, the results in the Appendices section suggest that the direction of this relationship remains the same. These polarizing effects suggest that when users go their own way, their autonomy is not as intact as for those that comply. This presents a potentially difficult trade-off between proactively steering users in a pro-social direction while at the same time respecting their autonomy, from users' own perspectives. In this regard, our qualitative analysis of users' comments indicated highly heterogeneous perceptions, ranging from positive opinions to feelings of reactance. Some users considered the nudge patterns to be a legitimate tool to maximize donations in a pro-social context. Other users perceived them as pressuring and manipulative. This polarization underscores the ethically complex landscape of nudging, illustrating that nudges often need to be judged on a case-by-case basis [50].

When users in our study were faced with the clear choice of either sticking with, or changing their donation, only a small minority ($\approx$ 3.4% of users) changed their donation amounts. Furthermore, this regretted donation amount does not seem to be particularly associated with any of the nudge combinations. While this is a positive sign, our findings do suggest that the decision to change the donation amount was associated with a higher threat to autonomy and sense of manipulation, see Figure 5.6. If users changed their donation amount at the end of the study, it can be argued that in those particular cases, these pro-social nudge combinations fail the practical notion of a "regret test" as outlined by Eyal [360]. We grant that regret is usually viewed as an emotion related to a decision-process [367], but in the related work we echoed the point that especially a sense of manipulation may not be apparent until after a given choice has been made [334]. While our findings on this point only relate to a very small portion of our users, there is still a clear connection between threats to autonomy/sense of manipulation and the decision to change the donation amount. Offending users in this sense, i.e. where they later regret their donation decision should arguable be avoided at high costs for online charities. It may be a good practice to verify a donation decision if a particular nudge was used in order to mitigate these ethical concerns.

It is worth noting that despite the overall increase in perceived threats to autonomy and sense of manipulation shown in the quantitative results, some participants also commented that the nudge patterns had a positive impact on their user experience as it facilitated their donation choices. Based on users' comments, this could be attributed to the fact that some nudge combinations, mainly the Fri.+Def.+Soc. seemed to alleviate their cognitive load by channeling their decision-making process, making it easier for them to make what they perceived to be an adequate donation.

Our results, overall, indicate that while some nudge patterns in a pro-social context can work to maximize donations, they come with hidden costs in terms of users' manipulation and threats to autonomy. As such, if pro-social nudges are used to increase donations, they should probably strike a balance between persuasion and perceived coercion. Hopefully, this can be accomplished by prioritizing more transparent, ethical designs that leaves the potential donor in a position where they, at least, do not regret their otherwise altruistic decision to donate towards a given online charity.

5.7.3 Limitations and future research

Future research could address several limitations of the present chapter. The first limitation of this study relates to the geographical location of participants, as they were all recruited in the UK. This could have influenced their reactions to the nudge patterns they were exposed to, skewing our results towards a specific cultural and economic background. Future research could mitigate this by including a more globally diverse participant pool [368], thereby offering a more nuanced understanding of the influence of digital nudge combinations on charitable giving across various cultural and socio–economic contexts.

Second, the design space that we used for investigating our nudge combinations is more or less infinite. That is, there is a great many ways to design various donation options, pre-selecting them (default), toggling the choice-related effort to reach some options (friction), and using various social design ques to guide behaviors (social). Our designs explored a very specific area on this landscape where digital nudges were combined to steer users towards donating the highest amount (20%) from a set of options. Variations of our designs could be further assessed in future research. One possible path worth exploring would be to separate the two factors of the design of digital nudges, and their direction. To illustrate, previous researchers have compared the acceptance of pro-social vs. pro-self nudge policies through several surveys [369, 318]. If our specific nudge combinations were pointed in the opposite direction (i.e. pro-self) and users' perspectives on threats to autonomy and sense of manipulation came out differently, this adds further nuance to the ethical discussion in the sense that it might not be the design of the nudge per se that triggers these concerns, but rather its direction. Untangling these two factors of nudge design vs. direction is a promising avenue of future research*.

Third, the monetary options that we used in our solution are very small (0-20 pence). Based on previous research it is not clear exactly how this may influence our findings. Small stakes may play a role in our inability to detect an effect from some of the nudge combinations, but a meta-analysis on the topic of stake-sizes suggests that, in the aggregate, small stakes should still provide a reasonable approximation of donation behavior [370]. Either way, high-powered online field experiments with larger stakes could further test the generalizability of our current findings.

Fourth, our evaluation of these nudge combinations does not allow us to rule out novelty effects. It could be argued that the efficacy of the nudge combinations con-

*While this is theoretically interesting, we do grant that charities have few practical incentives to steer online users towards donating as little as possible.

taining friction is associated to users not having encountered them before. However, various types of friction are very pervasive online. For example, when surveying online news outlets in 2019, researchers found that 69% of consent notices required 10 or more clicks to opt of all cookies [371]. This indicates that users do encounter quite extreme applications of design friction in online environments. Therefore, the effectiveness of friction nudges in our study should not solely be attributed to their novelty. While previous research has investigated the repeated impact of default nudges related to online charitable giving [342], future research could further investigate combinations of nudges in the same vein. Given the polarizing effects on users' perceptions of autonomy and manipulation that we detect in our study, it would be highly relevant to know if these potential effects have more long-term consequences on users' charitable giving behaviors.

5.8 Conclusion

With online giving taking up an increasing proportion of charitable giving, we investigated the impacts of pro-social nudge combinations on users' charitable giving behaviors, and their subsequent effects on users' ethical concerns towards these designs. We applied a design science research approach to evaluate a solution that systematically combined default, friction and social nudges. Our evaluation suggests that stacking 2-3 nudges did have an impact on donations and users' tendency to comply with the nudges, but once we account for interaction effects, friction appears to be the driving factor for increased donations. However, nudge combinations with friction in particular seemed to have polarizing effects on perceived threats to autonomy and the sense of manipulation on users that comply with these nudges and users that do not.

Finally, our qualitative analysis highlighted a complex and nuanced ethical landscape of nudging, underscoring the importance of considering individual differences and context-specific factors when these digital interventions are used in pro-social contexts.

6 Main Contributions, Future Research and Closing Remarks

With the passage of time, the psychology of people stays the same, but the tools and objects in the world change.

– Don Norman [372, p. 18]

Contents

6.1 Summary of Contributions

The chapters of the dissertation have explored the design space of digital choice architecture within several different application domains with the aim of addressing four over-arching questions related to the design of digital choice architecture and nudges. These questions were approached through four distinct but complementary research projects. To recap, the dissertation provided an overview of empirical studies in this design space together with three user studies using a design science research methodology. The main contributions of the dissertation will now be summarized.

The second chapter advanced the growing body of literature on digital nudging by providing a broad overview of empirical studies on the topic. Our review contained an expansive corpus of empirical work from several disciplines, which enabled us to provide novel insights for future research on the topic of digital nudging and choice architecture. The chapter contained a descriptive analysis of 73 papers, based on 109 studies containing 231 digital nudges that had been demonstrated and evaluated by previous researchers. Our results indicated a strong upward trend for research interest within the domain digital nudging. We then proceeded to highlight unexpected pockets of research attention across the designs, contexts, and evaluation methods related to digital nudging. We reasoned that interconnected and personalized choice architectures often differentiate digital from conventional or offline nudges. Based on our analysis, most of the research on digital nudging was still conducted in what we called disconnected and non-personalized choice architectures. About half of the research on digital nudging fell inside three main application contexts: privacy/security, e-commerce/marketing, and social media. The type of behavioral outcome that was measured tended to be an activation of a new behavior (rather than a cessation of an already established behavior) when nudges were applied in the literature. Few single studies measured both attitudinal and behavioral outcomes, but an increasing number of papers conducted several studies to better triangulate the outcomes of their nudging designs. We also found that the great majority of digital nudges were delivered through visual interfaces across mobile or desktop computers. Furthermore, several nudge combinations were very rarely investigated in the literature, such as friction coupled with reinforcement or warnings together with scarcity nudges. Based on the analyzed literature, we provided HCI scholars with nine future research questions that could facilitate our scientific understanding of this interdisciplinary research domain.

Chapter three was a novel attempt at tackling the important design challenge of increasing users 'informed consent' in the digital sphere. The problem that the third chapter addressed was based on the premise that service providers have ever more complex and sometimes disturbing agreements (ToS and PP) around when, how and why personal data is being collected and shared. Against this background, prior research suggested that a very small proportion of users are aware of this state of affairs. Because of this, they frequently commit what is known as the biggest lie on the Internet [72], i.e. users are unlikely to ever read the ToS and PP of a service provider before they use a given service. Through an online experiment and a descriptive field study we demonstrated how HCI and system designers could toggle the privacy awareness of users with digital nudging. One of our proposed designs showed roughly a 75% decrease in number of users who just agreed the ToS and PP without even viewing them. This significant reduction in users who committed the so-called biggest lie on the Internet was a promising first step. While the large-scale field study within the IKEA Place app demonstrated that these design changes could be used by large service providers, our experimental results showed that none of the proposed choice architectures inherently increased the recall of ToS and PP contents among users.

On average, clicking the ToS and PP did increase the user's recall of the contents in our experimental study, which was still an important signal.

Chapter four was focused on another pressing challenge of how to mitigate compulsive social media usage by design. While the Facebook newsfeed undoubtedly brings value to countless users it has also been described by prior researchers as one of the principle features for prolonging time on the platform itself [264], and this time can be associated with regret among users [51]. We deployed a field experiment where Facebook users were randomly assigned to different newsfeed diets, which, in turn, changed the default state of users' endless newsfeeds. Our results showed that both of the newsfeed diets did not statistically interact with users self-rated degree of compulsive use, suggesting these design interventions seemed to moderate the time spent on the platform for both more compulsive as well as less compulsive users. Our analysis suggested that the strict newsfeed diet did significantly decrease the time spent on the platform, but it came with negative consequences in terms of usability. When users were permitted to self-regulate their newsfeed diet the results still pointed to a decrease in time spent on the platform – without these steep costs to the user experience. Also, the exploratory thematic analysis showed that certain themes such as FOMO and Self-awareness emerged more regularly related to the experience of the users in the strict newsfeed diet.

The fifth chapter explored the effects of several pro-social nudge combinations on users' donation behaviors, and investigated ethical concerns from a user perspective: perceived threats to autonomy and sense of manipulation. We systematically combined default, friction and social nudges to understand their additive and interactive effects in a digital context where users could give away a real portion of their earnings.

Our evaluation suggested that when stacking two to three of these pro-social nudges, donations increased on average, but friction was the driving factor of this change. Yet, combinations of friction nudges had polarizing effects on users. Users who resisted combinations of friction nudges were much more likely to perceive threats to their autonomy, and experience a sense of manipulation when compared to users who complied with these nudges. Users' qualitative perspectives also demonstrated these polarizing effects, with reports of perceived pressure and manipulation by the nudges while some users reported that their choices were facilitated by these nudge combinations.

To conclude this summary, Table 6.1 on the next page maps the main contributions of the dissertation relative to the higher-level research questions that were outlined based on the literature review in second chapter.

Table 6.1: Empirical investigations and research contributions in the dissertation relative to the nine future research questions that were outlined in the second chapter.

	Future research questions that emerged from the literature review	Chapter
1:	Can the outcomes of a digital nudging be transposed to different classes of problems?	(3, 4, 5)
2:	Are the effects of digital nudges moderated by various cultural contexts?	
3:	How can the impact on users from interconnected and personalized choice architectures be better quantified and understood?	(4)
4:	How can effective personalized nudges be designed using minimal real-time contextual data?	
5:	What are the emergent ethical boundaries concerning personalized nudging?	
6:	Are the effects of digital nudges moderated by the delivery device or channel?	
7:	Is a given digital nudge as effective at stopping or decreasing a behavior as it is at initiating or increasing a behavior?	
8:	Is a given digital nudge as effective at changing repeated behaviors as it is at changing one-time behaviors?	
9:	How do digital nudges interact and through which mechanisms?	(5)

Future research questions are adopted from Bergram et al. [30]

Since the chapters on online privacy, social media, and online charitable giving tackled distinct problems - it can be argued that the complementary contribution from these empirical investigations in chapter three to five shine some degree of light on the first future question i.e. whether digital nudges can be transposed to different classes of problems. Chapters four and five also provided design demonstrations that take initial steps regarding two of the future research questions that we outlined in the literature review. Yet, Table 6.1 mainly underscores that there are several knowledge gaps left to fill for future researchers. At the time of writing, several HCI researchers have already picked up the mantle and contributed to various associated questions.

Examples of such contributions include the design and evaluation of digital nudges that foster deeper engagement for repeated study behaviors in the context of self-regulated learning among students [373], reviewing and further structuring the future research agenda on personalized digital nudging [374], or exploring what employees consider as ethically acceptable nudges in occupational health and wellbeing programs when the use of wearable devices is incentivized [375].

While chapters two to five in the dissertation outline their own local limitations and paths for future research, there are still several promising avenues where future research directions can be articulated.

6.2 Future Research Directions

These future research avenues concern emergent delivery channels related to digital nudging and their implications, together with the challenge of mapping privacy outcomes to any digital choice architecture.

6.2.1 Digital nudging across different delivery channels

In chapter two, the delivery channels that we encountered in the literature was almost exclusively visual. That is, digital nudges were designed and evaluated using various kinds of screens. One the one hand, this makes sense since other channels, for example voice user interfaces (VUIs) are a quite recent phenomena [376]. On the other hand, these channels are now ubiquitous with companies like Amazon announcing that they have so far sold more than half-a-billion Alexa-enabled devices [377]. Other prominent examples of VUIs would include Google's Assistant, Apple's Siri, or Microsoft's Cortana.

Channels that are not visual present a whole set of challenges that will need to be addressed by future researchers and practitioners. There are also ethical design challenges that arise here around transparency. Empirically inspecting a users' interactions with a graphical UI can give a lot of meaningful information around how choices are presented and structured. VUIs, on the other hand, make a potential choice architecture extremely opaque. Given the transient nature of voice interaction, potential alternatives routes for the user become very hidden. That is, inspecting how certain choice alternatives are being reinforced or weighted is easier to do in the visual realm, than through audio channels. Transparency therefore becomes an issue as soon as secondary data is involved. From our current perspective there also seems to be conceptual challenges that arise once we abandon visual delivery channels for digital nudges. While it is interesting to consider the moderating effects in terms of nudging from different delivery channels, lets consider the comparison of specific digital nudges across visual and audio channels.

For example, defaults can be represented through many visual design elements, such as a pre-selected button, a pre-written input field, a highlighted alternative in a drop-down list, default settings etc., but how can defaults be instantiated in VUIs? There will of course be default rules and settings for voice interactions that HCI and system designers will have to set, for example: how many alternatives should be provided to the user when the user has not specified that number themselves. Maybe the default rule should be to always prompt the user to clarify how many alternatives they want? But if we do not consider default rules, but the finer question of how to represent a default outcome (like a pre-selected button) for the user, the answer seems far away from clear. Similar questions can of course be put for a variety of nudge patterns. This translation work has already begun for design principles and heuristics in general [376], but to our knowledge this has not yet happened for the design process of digital nudges. Both novel design approaches and an inclination to survey the work of other associated research domains will be needed guide the design of digital nudges across other delivery channels. However, given the emerging ubiquity of other channels of interaction (especially VUI), this seems like an important avenue for future research.

6.2.2 Developing more meaningful online privacy controls

In chapter three we echoed that most users are not that well informed regarding what they agree to in online settings [72]. Achieving a good level of online privacy awareness for users is a difficult challenge because it suffers from a mapping problem. That is, privacy as an outcome is not well defined, and designing a digital choice architecture where various design elements are mapped towards this outcome is thus challenging. Privacy obviously must include some degree of control over some information about the user herself [378]. However, giving the user an overview of what this information might entail and the proper digital choice architecture for controlling its dissemination – that is where the true design challenge lies. As we saw in chapter three, the current standard around online privacy harks back to a legal framework known as notice-and-consent, in line with Obar and Oeldorf-Hirsch [216], we referred to this standard as "quick-join". Other researchers have argued that given this standard, online privacy is a futile effort because of a transparency paradox [379]. That is, transparency would involve giving the user a complete overview of what personal information they might be sharing with an online service provider. However, if this complete overview is provided, the user is very unlikely to understand it, let alone read it [379]. I believe that researchers and practitioners can eventually improve if not overcome the problems related to this transparency paradox. While still in its early stages, digital nudging research has already been performed in this vein.

We previously referenced Schöbel et al. [121], who studied users' attitudes towards several digital nudging patterns (default, feedback, social, warnings, etc.) in the context of online privacy. Their study revealed that their sample of Slack users preferred to be nudged towards privacy by especially defaults and red/green warning graphics. Understanding these attitudes towards online privacy is an important step, but more studies also need to consider how to traverse the attitude–behavior gap concerning online privacy. While chapter three in this dissertation is an earnest attempt to do just that, there is still a lot of room for improvements. Future research will be able to leverage new interactive technologies such as intelligent agents representing user interests, or the usage of privacy-preserving analytics algorithms among data holders [380]. Hopefully, this will open up many novel opportunities for tackling this important challenge.

6.3 Closing Remarks

Interactive technologies are going to push societies into a corner where it becomes increasingly consequential to set certain boundaries around what kind of digital persuasion that is deemed permissible. Digital nudging is still not a panacea to the design of behavior change interventions in general – it is just one of several applicable approaches. Yet, my present perspective is that there in no door on offer in the 21st century that says 'no digital nudging here'. Digital choice architecture is inevitable, and great care should taken when HCI and system designers try to shape online decision environments for prospective users. On this topic, the dissertation has provided an extensive review of previous empirical studies, and three user studies that tackled both theoretical and practical problems related to the application domains of online privacy, social media, and online charitable giving.

Nowadays, we live in a world where users are often unwittingly enrolled (depending on which apps they use) in split tests where their choices are used to determine how a particular digital choice architecture should be updated. The future remains uncertain, and it is an open-ended question whether public and private organizations will be able to design and maintain digital choice architectures that make all of us better of, *as judged by ourselves.* In closing, my hope is that this dissertation advances our design scientific knowledge related to this consequential area of research and everyday life.

7 References

The chapters in this dissertation were based on a collection of articles. The references are therefore indexed on each chapter. Sources may appear across several of these chapters. As a rule, the references to these sources will be listed under the chapter where they first appeared in the dissertation.

1. Introduction and Definition of the Design Space

[1] Richard H Thaler and Cass R Sunstein. *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT, USA: Yale University Press, 2008.

[2] Agnieszka Kitkowska. "The Hows and Whys of Dark Patterns: Categorizations and Privacy". In: *Human Factors in Privacy Research*. Ed. by Nina Gerber, Alina Stöver, and Karola Marky. 2023, pp. 173–198.

[3] ISO Central Secretary. *Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems*. Standard ISO 9241-210:2010. International Organization for Standardization, 2010. URL: `https://www.iso.org/standard/52075.html`.

[4] *Pew Research Center.* The Virtues and Downsides of Online Dating. Website, Accessed on: 30/06/2020. Available at: `https://www.pewresearch.org/internet/2020/02/06/the-virtues-and-downsides-of-online-dating/`, 2020.

[5] *European Banking Federation.* Structure and Economic Contribution of the Banking Sector. Website, Accessed on: 30/06/2020. Available at: `https://www.ebf.eu/facts-and-figures/structure-and-economic-contribution-of-the-banking-sector/`, 2020.

[6] Anthony Jameson et al. "Choice Architecture for Human-Computer Interaction". In: *Interaction* 7.1-2 (2013), pp. 1–235.

[7] Herbert A Simon. "A behavioral model of rational choice". In: *The quarterly journal of economics* 69.1 (1955), pp. 99–118.

[8] Herbert A Simon. *Models of Man; social and rational*. Wiley, 1957.

[9] Amos Tversky and Daniel Kahneman. "Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty." In: *science* 185.4157 (1974), pp. 1124–1131.

[10] Esther-Mirjam Sent. "Rationality and bounded rationality: You can't have one without the other". In: *The European Journal of the History of Economic Thought* 25.6 (2018), pp. 1370–1386.

[11] Thomas Mussweiler and Fritz Strack. "Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model". In: *Journal of Experimental Social Psychology* 35.2 (1999), pp. 136–164.

[12] Dan Ariely and Simon Jones. *Predictably irrational*. Harper-Collins, New York, 2008.

[13] Solomon E Asch. "Opinions and social pressure". In: *Scientific American* 193.5 (1955), pp. 31–35.

[14] Andreas Mojzisch and Kristine Krug. "Cells, circuits, and choices: Social influences on perceptual decision making". In: *Cognitive, Affective, & Behavioral Neuroscience* 8 (2008), pp. 498–508.

[15] Kazuo Mori and Miho Arai. "No need to fake it: Reproduction of the Asch experiment without confederates". In: *International Journal of Psychology* 45.5 (2010), pp. 390–397.

[16] Aiko Hanayama, Kazuo Mori, et al. "Conformity of six-year-old children in the Asch experiment without using confederates". In: *Psychology* 2.7 (2011), pp. 661–664.

[17] Abhijit V Banerjee. "A simple model of herd behavior". In: *The quarterly journal of economics* 107.3 (1992), pp. 797–817.

[18] Ryan O Murphy and Kurt A Ackermann. "Social value orientation: Theoretical and measurement issues in the study of social preferences". In: *Personality and Social Psychology Review* 18.1 (2014), pp. 13–41.

[19] Stephen Worchel, Jerry Lee, and Akanbi Adewole. "Effects of supply and demand on ratings of object value." In: *Journal of personality and social psychology* 32.5 (1975), p. 906.

[20] Daniel Kahneman and Amos Tversky. "Prospect theory: An analysis of decision under risk". In: *Econometrica* 47.2 (1979), pp. 263–292.

[21] Daniel Kahneman, Jack L Knetsch, and Richard H Thaler. "Experimental tests of the endowment effect and the Coase theorem". In: *Journal of political Economy* 98.6 (1990), pp. 1325–1348.

[22] Amos Tversky and Daniel Kahneman. "The framing of decisions and the psychology of choice". In: *Science* 211.4481 (1981), pp. 453–458.

[23] Arvind Narayanan, Arunesh Mathur, Marshini Chetty, and Mihir Kshirsagar. "Dark Patterns: Past, Present, and Future". In: *Communication of the ACM* 63.9 (2020), pp. 42–47.

[24] Richard H Thaler. *Misbehaving: The making of behavioral economics*. Penguin Random House UK, 2015.

[25] Eric J Johnson et al. "Beyond nudges: Tools of a choice architecture". In: *Marketing Letters* 23.2 (2012), pp. 487–504.

[26] Graham Lawton. *Nudge: You're being manipulated – have you noticed?* Website. Retrieved December 9, 2023 from: `https://www.newscientist.com/article/mg21829221-700-nudge-youre-being-manipulated-have-you-noticed/`. 2013.

[27] Markus Weinmann, Christoph Schneider, and Jan vom Brocke. "Digital nudging". In: *Business & Information Systems Engineering* 58.6 (2016), pp. 433–436.

[28] Natasha D Schüll. "Data for life: Wearable technology and the design of self-care". In: *BioSocieties* 11 (2016), pp. 317–333.

[29] Eric Horvitz and Deirdre Mulligan. "Data, privacy, and the greater good". In: *Science* 349.6245 (2015), pp. 253–255.

[30] Kristoffer Bergram, Marija Djokovic, Valéry Bezençon, and Adrian Holzer. "The digital landscape of nudging: A systematic literature review of empirical research on digital nudges". In: *Proceedings of the 2022 CHI conference on human factors in computing systems*. New Orleans, Louisiana, USA: ACM, 2022, pp. 1–16.

[31] Cass R Sunstein. "The ethics of nudging". In: *Yale Journal on Regulation* 32.2 (2015), pp. 413–450.

[32] Charlie Pinder, Jo Vermeulen, Benjamin R Cowan, and Russell Beale. "Digital behaviour change interventions to break and form habits". In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 25.3 (2018), pp. 1–66.

[33] Brian J Fogg. "Persuasive computers: perspectives and research directions". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., 1998, pp. 225–232.

[34] BJ Fogg. *Persuasive Technology: Using computers to change what we think and do.* Morgan Kaufmann Publishers, San Francisco, 2003.

[35] Harri Oinas-Kukkonen and Marja Harjumaa. “A systematic framework for designing and evaluating persuasive systems”. In: *Persuasive Technology: Third International Conference, PERSUASIVE’08, Oulu, Finland, June 4-6, 2008.* Springer, 2008, pp. 164–176.

[36] Harri Oinas-Kukkonen and Marja Harjumaa. “Persuasive systems design: Key issues, process model, and system features”. In: *Communications of the Association for Information Systems* 24.28 (2009), pp. 485–500.

[37] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. “The Dark (Patterns) Side of UX Design”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* CHI ’18. Montreal QC, Canada: ACM, 2018, pp. 1–14.

[38] Harry Brignull. *Dark Patterns: Deception vs. Honesty in UI Design.* Website, A LIST APRART. Retrieved October 29, 2023 from: `https://alistapart.com/article/dark-patterns-deception-vs-honesty-in-ui-design/`. 2011.

[39] Gregory Conti and Edward Sobiesk. “Malicious interface design: exploiting the user”. In: *Proceedings of the 19th international conference on World wide web, WWW’10.* 2010, pp. 271–280.

[40] Arunesh Mathur et al. “Dark patterns at scale: Findings from a crawl of 11K shopping websites”. In: *Proceedings of the ACM on Human-Computer Interaction* 3.CSCW (2019), pp. 1–32.

[41] José P Zagal, Staffan Björk, and Chris Lewis. “Dark patterns in the design of games”. In: *Foundations of Digital Games, FDG’2013.* May 14-17, Chania, Greec, 2013.

[42] Jamie Luguri and Lior Jacob Strahilevitz. “Shining a light on dark patterns”. In: *Journal of Legal Analysis* 13.1 (2021), pp. 43–109.

[43] Jen Caltrider, Misha Rykov, and MacDonald Zoë. *It’s Official: Cars Are the Worst Product Category We Have Ever Reviewed for Privacy.* Website, Accessed on: 30/09/2023. Available at: `https://foundation.mozilla.org/en/`

privacynotincluded/articles/its-official-cars-are-the-worst-product-category-we-have-ever-reviewed-for-privacy/, 2023.

[44] Steve Rathje, Jay J Van Bavel, and Sander Van Der Linden. "Out-group animosity drives engagement on social media". In: *Proceedings of the National Academy of Sciences* 118.26 (2021), e2024292118.

[45] William J Brady, Julian A Wills, John T Jost, Joshua A Tucker, and Jay J Van Bavel. "Emotion shapes the diffusion of moralized content in social networks". In: *Proceedings of the National Academy of Sciences* 114.28 (2017), pp. 7313–7318.

[46] Abhijnan Chakraborty, Nuno Mota, Asia J Biega, Krishna P Gummadi, and Hoda Heidari. "On the impact of choice architectures on inequality in online donation platforms". In: *Proceedings of WWW '19*. San Francisco, CA, USA: ACM, 2019, pp. 2623–2629.

[47] Robert Epstein, Ronald E Robertson, David Lazer, and Christo Wilson. "Suppressing the search engine manipulation effect (SEME)". In: *Proceedings of the ACM on Human-Computer Interaction* 1.CSCW (2017), pp. 1–22.

[48] Fabian Haak and Philipp Schaer. "Auditing search query suggestion bias through recursive algorithm interrogation". In: *Proceedings of the 14th ACM Web Science Conference 2022*. Bercelona, Spain: ACM, 2022, pp. 219–227.

[49] Thomas H Davenport and John C Beck. *The Attention Economy*. Harvard Buniness School Press, 2001.

[50] Andreas T Schmidt and Bart Engelen. "The ethics of nudging: An overview". In: *Philosophy compass* 15.4 (2020), e12658.

[51] Thomas Mildner and Gian-Luca Savino. "Ethical User Interfaces: Exploring the Effects of Dark Patterns on Facebook". In: *Proceedings of the CHI conference on human factors in computing systems*. CHI EA '21. Yokohama, Japan: ACM, 2021, pp. 1–7.

[52] Kristoffer Bergram, Valéry Bezençon, Paul Maingot, Tony Gjerlufsen, and Adrian Holzer. "Digital Nudges for Privacy Awareness: From consent to informed consent?" In: *Proceedings of ECIS'20*. Marrakech, Morocco: AIS, 2020, pp. 1–16.

[53] Aditya Purohit, Kristoffer Bergram, Louis Barclay, Valéry Bezençon, and Adrian Holzer. "Starving the Newsfeed for Social Media Detox: Effects of Strict and Self-regulated Facebook Newsfeed Diets". In: *Proceedings of the 2023 CHI conference on human factors in computing systems*. Hamburg, Germany: ACM, 2023, pp. 1–16.

[54] Kristoffer Bergram, Abdessalam Ouaazki, Manon Berney, Valéry Bezençon, and Adrian Holzer. *Stacking Nudges for Online Charitable Giving: An exploratory analysis of pro-social nudge combinations and their polarizing effects from a user's perspective*. Working paper. University of Neuchâtel, 2023.

2. Digital Landscape of Nudging and Choice Architecture

[55] William Eggers et al. *Government Trends 2020 report*. Tech. rep. Deloitte, 2020.

[56] Cass R Sunstein. "Nudges, agency, and abstraction: A reply to critics". In: *Review of Philosophy and Psychology* 6.3 (2015), pp. 511–529.

[57] Richard H Thaler, Cass R Sunstein, and John P Balz. "Choice architecture". In: *The behavioral foundations of public policy*. Princeton, NJ, USA: Princeton University Press, 2013, pp. 428–439.

[58] Pelle Guldborg Hansen. "The definition of nudge and libertarian paternalism: Does the hand fit the glove?" In: *European Journal of Risk Regulation* 7.1 (2016), pp. 155–174.

[59] Christian Meske and Tobias Potthoff. "The DINU-model–a process model for the design of nudges". In: *Proceedings of ECIS'17*. Guimarães, Portugal: AIS, 2017, pp. 1–11.

[60] Tim-Benjamin Lembcke, Nils Engelbrecht, Alfred Benedikt Brendel, Bernd Herrenkind, and Lutz M Kolbe. "Towards a Unified Understanding of Digital Nudging by Addressing its Analog Roots". In: *Proceedings of PACIS'19*. Xi'an, China: AIS, 2019, pp. 1–14.

[61] Adrien Barton and Till Grüne-Yanoff. "From libertarian paternalism to nudging—and beyond". In: *Review of Philosophy and psychology* 6.3 (2015), pp. 341–359.

[62] Leonhard K Lades and Liam Delaney. "Nudge FORGOOD". In: *Behavioural Public Policy* 6.1 (2020), pp. 75–94.

[63] Daniel M Hausman and Brynn Welch. "Debate: To nudge or not to nudge". In: *Journal of Political Philosophy* 18.1 (2010), pp. 123–136.

[64] Richard H. Thaler. "Nudge, not sludge". In: *Science* 631.6401 (2018), p. 431.

[65] Jaejeung Kim, Joonyoung Park, Hyunsoo Lee, Minsam Ko, and Uichin Lee. "LocknType: Lockout task intervention for discouraging smartphone app use". In: *Proceedings of the 2019 CHI conference on human factors in computing systems*. Glasgow, Scotland, UK: ACM, 2019, pp. 1–12.

[66] Christin Hoffmann and Kirsten Thommes. "Can digital feedback increase employee performance and energy efficiency in firms? Evidence from a field experiment". In: *Journal of Economic Behavior & Organization* 180 (2020), pp. 49–65.

[67] Annamina Rieder, U Yeliz Eseryel, Christiane Lehrer, and Reinhard Jung. "Why Users Comply with Wearables: The Role of Contextual Self-Efficacy in Behavioral Change". In: *International Journal of Human-Computer Interaction* 37.3 (2021), pp. 281–294.

[68] Robert Epstein and Ronald E Robertson. "The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections". In: *Proceedings of the National Academy of Sciences* 112.33 (2015), E4512–E4521.

[69] Soroush Vosoughi, Deb Roy, and Sinan Aral. "The spread of true and false news online". In: *Science* 359.6380 (2018), pp. 1146–1151.

[70] Manoel Horta Ribeiro, Raphael Ottoni, Robert West, Virgílio AF Almeida, and Wagner Meira Jr. "Auditing radicalization pathways on YouTube". In: *Proceedings of FAT* '20*. Barcelona, Spain: ACM, 2020, pp. 131–141.

[71] Mark Alfano, Amir Ebrahimi Fard, J Adam Carter, Peter Clutton, and Colin Klein. "Technologically scaffolded atypical cognition: The case of YouTube's recommender system". In: *Synthese* 199. (2021), pp. 835–858.

[72] Jonathan A Obar and Anne Oeldorf-Hirsch. "The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services". In: *Information, Communication & Society* 23.1 (2020), pp. 128–147.

[73] Dennis Hummel and Alexander Maedche. "How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies". In: *Journal of Behavioral and Experimental Economics* 80 (2019), pp. 47–58.

[74] Ana Caraban, Evangelos Karapanos, Daniel Gonçalves, and Pedro Campos. "23 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow, Scotland, UK: ACM, 2019, pp. 1–15.

[75] T. Mirsch, C. Lehrer, and R. Jung. "Digital nudging: Altering user behavior in digital environments". In: *Proceedings of International Conference on Wirtschaftsinformatik*. St.Gallen, Switzerland: AIS, 2017, pp. 634–648.

[76] Agnieszka Kitkowska, Yefim Shulman, Leonardo A Martucci, and Erik Wästlund. "Psychological effects and their role in online privacy interactions: A review". In: *IEEE Access* 8 (2020), pp. 21236–21260.

[77] Mathias Jesse and Dietmar Jannach. "Digital nudging with recommender systems: Survey and future directions". In: *Computers in Human Behavior Reports* 3 (2021), pp. 1–14.

[78] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, Prisma Group, et al. "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement". In: *PLoS Medicine* 6.7 (2009), e1000097.

[79] Alessandro Liberati et al. "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration". In: *Journal of clinical epidemiology* 62.10 (2009), e1–e34.

[80] Jan vom Brocke et al. "Reconstructing the giant: On the importance of rigour in documenting the literature search process". In: *Proceedings of ECIS'09*. Verona, Italy: AIS, 2009, pp. 1–12.

[81] Juho Hamari, Jonna Koivisto, and Harri Sarsa. "Does gamification work?– A literature review of empirical studies on gamification". In: *Proceedings of HICSS'14*. Waikoloa, HI, USA: IEEE, 2014, pp. 3025–3034.

[82] Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub, and Thorsten Holz. "(Un)informed consent: Studying gdpr consent notices in the field". In: *Proceedings of CCS'19*. London, UK: ACM, 2019, pp. 973–990.

[83] Vincent W-S Tseng, Matthew L Lee, Laurent Denoue, and Daniel Avrahami. “Overcoming distractions during transitions from break to work using a conversational website-blocking system”. In: *Proceedings of CHI’19*. Glasgow, Scotland, UK: ACM, 2019, pp. 1–13.

[84] Tianshu Sun, Siva Viswanathan, Ni Huang, and Elena Zheleva. “Designing Promotional Incentive to Embrace Social Sharing: Evidence from Field and Lab Experiments”. In: *Proceedings of ICIS’19*. Munich, Germany: AIS, 2019, pp. 1–17.

[85] Aarathi Prasad and Asia Quinones. “Digital Overload Warnings-“The Right Amount of Shame”?” In: *Proceedings of International Conference on Human-Computer Interaction*. Copenhagen, Denmark: Springer, 2020, pp. 117–134.

[86] Martin Kretzer and Alexander Maedche. “Designing social nudges for enterprise recommendation agents: An investigation in the business intelligence systems context”. In: *Journal of the Association for Information Systems* 19.12 (2018), pp. 1145–1186.

[87] Olga Abramova. “What’s in a name: Examining discrimination on sharing platforms and its possible remedies”. In: *Proceedings of ECIS’20*. Marrakech, Morocco: AIS, 2020, pp. 1–17.

[88] Antino Kim and Alan R Dennis. “Says who? The effects of presentation format and source rating on fake news in social media.” In: *MIS Quarterly* 43.3 (2019).

[89] Thomas M Connolly, Elizabeth A Boyle, Ewan MacArthur, Thomas Hainey, and James M Boyle. “A systematic literature review of empirical evidence on computer games and serious games”. In: *Computers & education* 59.2 (2012), pp. 661–686.

[90] Jane Webster and Richard T Watson. “Analyzing the past to prepare for the future: Writing a literature review”. In: *MIS quarterly* 26.2 (2002), pp. 13–23.

[91] Sally Fincher et al. “Perspectives on HCI patterns: concepts and tools”. In: *Proceedings of CHI’03 Extended Abstracts*. Fort Lauderdale, FL, USA: ACM, 2003, pp. 1044–1045.

[92] Cass R Sunstein. “Nudging: a very short guide”. In: *Journal of Consumer Policy* 37.4 (2014), pp. 583–588.

[93] Noah J Goldstein, Robert B Cialdini, and Vladas Griskevicius. "A room with a viewpoint: Using social norms to motivate environmental conservation in hotels". In: *Journal of consumer Research* 35.3 (2008), pp. 472–482.

[94] Blake M DiCosola and Gina Neff. "Using Social Comparisons to Facilitate Healthier Choices in Online Grocery Shopping Contexts". In: *Proceedings of CHI'20 Extended Abstracts*. Honolulu, HI, USA: ACM, 2020, pp. 1–8.

[95] Andre Calero Valdez, Martina Ziefle, and Michael Sedlmair. "Priming and anchoring effects in visualization". In: *IEEE transactions on visualization and computer graphics* 24.1 (2017), pp. 584–594.

[96] Daniel M Oppenheimer, Robyn A LeBoeuf, and Noel T Brewer. "Anchors aweigh: A demonstration of cross-modality anchoring and magnitude priming". In: *Cognition* 106.1 (2008), pp. 13–26.

[97] Alan R Dennis, Lingyao Yuan, Xuan Feng, Eric Webb, and Christine J Hsieh. "Digital Nudging: Numeric and Semantic Priming in E-Commerce". In: *Journal of Management Information Systems* 37.1 (2020), pp. 39–65.

[98] Michael G Brown, James J Schiltz, Holly Derry, and Caitlin Holman. "Implementing online personalized social comparison nudges in a web-enabled coaching system". In: *The Internet and Higher Education* 43.1 (2019), p. 100691.

[99] Mark A Andor, Andreas Gerster, and Stephan Sommer. "Consumer inattention, heuristic thinking and the role of energy labels". In: *The Energy Journal* 41.1 (2020), pp. 83–112.

[100] Henner Gimpel, Sebastian Heger, Julia Kasper, and Ricarda Schäfer. "The Power of Related Articles–Improving Fake News Detection on Social Media Platforms". In: *Proceedings of HCISS'20*. Maui, HI, USA: ScholarSpace, 2020, pp. 6063–6072.

[101] Ashley M Fox, Edmund C Stazyk, and Wenhui Feng. "Administrative easing: Rule reduction and medicaid enrollment". In: *Public Administration Review* 80.1 (2020), pp. 104–117.

[102] Fabian Okeke, Michael Sobolev, Nicola Dell, and Deborah Estrin. "Good vibrations: can a digital nudge reduce digital overload?" In: *Proceedings of MobileHCI'18*. Barcelona, Spain: ACM, 2018, pp. 1–12.

[103] Christina L Brown and Aradhna Krishna. "The skeptical shopper: A metacognitive account for the effects of default options on choice". In: *Journal of consumer research* 31.3 (2004), pp. 529–539.

[104] Eric J Johnson and Daniel Goldstein. "Do defaults save lives?" In: *Science* 302.5649 (2003), pp. 1338–1339.

[105] Moritz Ingendahl, Dennis Hummel, Alexander Maedche, and Tobias Vogel. "Who can be nudged? Examining nudging effectiveness in the context of need for cognition and need for uniqueness". In: *Journal of Consumer Behaviour* 20 (2020), pp. 324–336.

[106] Dominik Jung and Christof Weinhardt. "Robo-advisors and financial decision inertia: How choice architecture helps to reduce inertia in financial planning tools". In: *Proceedings of ICIS'18*. San Francisco, CA, USA: AIS, 2018, pp. 1–17.

[107] Christian von der Weth, Ashraf Abdul, Shaojing Fan, and Mohan Kankanhalli. "Helping Users Tackle Algorithmic Threats on Social Media: A Multimedia Research Agenda". In: *Proceedings of MM'20*. Seattle, WA, USA: ACM, 2020, pp. 4425–4434.

[108] Robert B Cialdini. *Influence*. Boston, MA, USA: Pearson Education, 2001.

[109] Timm Teubner and Antje Graul. "Only one room left! How scarcity cues affect booking intentions on hospitality platforms". In: *Electronic Commerce Research and Applications* 39 (2020), p. 100910.

[110] Christoph Schneider, Markus Weinmann, and Jan Vom Brocke. "Digital nudging: guiding online user choices through interface design". In: *Communications of the ACM* 61.7 (2018), pp. 67–73.

[111] Michael Wessel, Martin Adam, and Alexander Benlian. "The impact of sold-out early birds on option selection in reward-based crowdfunding". In: *Decision Support Systems* 117 (2019), pp. 48–61.

[112] Robert B Cialdini and Noah J Goldstein. "Social influence: Compliance and conformity". In: *Annu. Rev. Psychol.* 55 (2004), pp. 591–621.

[113] Tobias Kroll, Ute Paukstadt, Kseniya Kreidermann, and Milad Mirbabaie. "Nudging people to save energy in smart homes with social norms and self-commitment". In: *Proceedings of ECIS'19*. Stockholm and Uppsala, Sweden: AIS, 2019, pp. 1–10.

[114] Armenak Antinyan, Gergely Horváth, and Mofei Jia. "Curbing the consumption of positional goods: Behavioral interventions versus taxation". In: *Journal of Economic Behavior & Organization* 179 (2020), pp. 1–21.

[115] David Schneider, Johannes Klumpe, Martin Adam, and Alexander Benlian. "Nudging users into digital service solutions". In: *Electronic Markets* 30.4 (2019), pp. 863–881.

[116] Christoph Buck, Simone Burster, and Torsten Eymann. "An Experiment Series on App Information Privacy Concerns". In: *Proceedings of ECIS'18*. Portsmouth, UK: AIS, 2018, pp. 1–18.

[117] Rohan Hundia, Aditya Vijayvargiya, and Aaron Quigley. "Closed loop feedback nudges using nerve stimulation". In: *Proceedings of SIGCHI on EICS'20*. Honolulu, HI, USA: ACM, 2020, pp. 1–6.

[118] Avijit Sengupta, Anol Bhattacherjee, and Kaushik Dutta. "Information Technology Interventions in Cardiac Rehabilitation: A Theory Driven Approach". In: *Proceedings of ICIS'20*. Hyderabad, India: AIS, 2020, pp. 1–17.

[119] Adam Beddoe et al. "Disruptabottle: Encouraging Hydration with an Overflowing Bottle". In: *Proceedings of CHI'20 Extended Abstracts*. Honolulu, HI, USA: ACM, 2020, pp. 1–7.

[120] Maxim van Oldenbeek, Till J Winkler, Julie Buhl-Wiggers, and Daniel Hardt. "Nudging in Blended Learning: Evaluation of Email-based progress feedback in a Flipped-Classroom Information Systems Course". In: *Proceedings of ECIS'19*. Stockholm and Uppsala, Sweden: AIS, 2019, pp. 1–15.

[121] Sofia Schöbel, Torben Barev, Andreas Janson, Felix Hupfeld, and Jan Marco Leimeister. "Understanding User Preferences of Digital Privacy Nudges–A Best-Worst Scaling Approach". In: *Proceedings of HICSS'20*. Maui, HI, USA: ScholarSpace, 2020, pp. 3918–3927.

[122] Brian J Fogg and Jason Hreha. "Behavior wizard: A method for matching target behaviors with solutions". In: *Proceedings of PERSUASIVE'10*. Copenhagen, Denmark: Springer, 2010, pp. 117–131.

[123] Raphael Schilling, Stephan Aier, and Robert Winter. "Designing an Artifact for Informal Control in Enterprise Architecture Management". In: *Proceedings of ICIS'19*. Munich, Germany: AIS, 2019, pp. 1–17.

[124] Antonio Gualberto Pereira and Luís Eduardo Afonso. "Automatic enrollment and employer match: an experiment with the choice of pension plans". In: *Revista de Gestão* 27.3 (2019), pp. 281–299.

[125] Guillaume Saint-Jacques, Amir Sepehri, Nicole Li, and Igor Perisic. *Building inclusive products through A/B testing*. Website. Retrieved March 9, 2021 from: `https://engineering.linkedin.com/blog/2020/building-inclusive-products-through-a-b-testing`. 2020.

[126] Adam DI Kramer, Jamie E Guillory, and Jeffrey T Hancock. "Experimental evidence of massive-scale emotional contagion through social networks". In: *Proceedings of the National Academy of Sciences* 111.24 (2014), pp. 8788–8790.

[127] Alex Hern. *Why Google has 200m reasons to put engineers over designers*. Website. Retrieved April 5, 2021 from: `https://www.theguardian.com/technology/2014/feb/05/why-google-engineers-designers`. 2014.

[128] Michel Tuan Pham. "The seven sins of consumer psychology". In: *Journal of Consumer Psychology* 23.4 (2013), pp. 411–423.

[129] Geert Hofstede. *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Thousand Oaks, CA, USA: Sage publications, 2001.

[130] Daphna Oyserman. "High power, low power, and equality: Culture beyond individualism and collectivism". In: *Journal of Consumer Psychology* 16.4 (2006), pp. 352–356.

[131] Michele J Gelfand, Lisa H Nishii, and Jana L Raver. "On the nature and importance of cultural tightness-looseness". In: *Journal of applied psychology* 91.6 (2006), pp. 1225–1244.

[132] Dominic DiFranzo and Kristine Gloria-Garcia. "Filter bubbles and fake news". In: *XRDS: Crossroads, The ACM Magazine for Students* 23.3 (2017), pp. 32–35.

[133] Paul Resnick, R Kelly Garrett, Travis Kriplean, Sean A Munson, and Natalie Jomini Stroud. "Bursting your (filter) bubble: strategies for promoting diverse exposure". In: *Proceedings of CSCW'13*. San Antonio, TX, USA: ACM, 2013, pp. 95–100.

[134] Brent Kitchens, Steven L Johnson, and Peter Gray. "Understanding Echo Chambers and Filter Bubbles: The Impact of Social Media on Diversification and Partisan Shifts in News Consumption". In: *MIS Quarterly* 44.4 (2020), pp. 1619–1650.

[135] Aditya Kumar Purohit and Adrian Holzer. "Functional digital nudges: Identifying optimal timing for effective behavior change". In: *Proceedings of CHI'19 Extended Abstracts*. Glasgow, Scotland, UK: ACM, 2019, pp. 1–6.

[136] Robert Bixler and Sidney D'Mello. "Automatic gaze-based user-independent detection of mind wandering during computerized reading". In: *User Modeling and User-Adapted Interaction* 26.1 (2016), pp. 33–68.

[137] Yohan Chon, Elmurod Talipov, Hyojeong Shin, and Hojung Cha. "Mobility prediction-based smartphone energy optimization for everyday location monitoring". In: *Proceedings of SenSys '11*. Seattle, WA, USA: ACM, 2011, pp. 82–95.

[138] Iñigo Torres Uribe Echebarría, Syed Anas Imtiaz, Mingxu Peng, and Esther Rodriguez-Villegas. "Monitoring smoking behaviour using a wearable acoustic sensor". In: *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Jeju, South Korea: IEEE, 2017, pp. 4459–4462.

[139] Sebastien Richoz, Lin Wang, Philip Birch, and Daniel Roggen. "Transportation mode recognition fusing wearable motion, sound, and vision sensors". In: *IEEE Sensors Journal* 20.16 (2020), pp. 9314–9328.

[140] Evangelia Anagnostopoulou et al. "From mobility patterns to behavioural change: leveraging travel behaviour and personality profiles to nudge for sustainable transportation". In: *Journal of Intelligent Information Systems* 54. (2020), pp. 1–22.

[141] Pelle Guldborg Hansen and Andreas Maaløe Jespersen. "Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy". In: *European Journal of Risk Regulation* 4.1 (2013), pp. 3–28.

[142] Eyal Peer et al. "Nudge me right: Personalizing online security nudges to people's decision-making styles". In: *Computers in Human Behavior* 109 (2020), p. 106347.

[143] Johnny Rodgers and Lyn Bartram. "Exploring ambient and artistic visualization for residential energy use feedback". In: *IEEE transactions on visualization and computer graphics* 17.12 (2011), pp. 2489–2497.

[144] Aleksander Gisvold and Esben Aarseth. *Using Gamification and Tangible User Interfaces to Treat Asthmatic Children.* Tech. rep. Institutt for datateknikk og informasjonsvitenskap, 2014.

[145] Adrian Holzer et al. "Gamifying knowledge sharing in humanitarian organisations: a design science journey". In: *European Journal of Information Systems* 29.2 (2020), pp. 153–171.

[146] Riccardo Miotto, Matteo Danieletto, Jerome R Scelza, Brian A Kidd, and Joel T Dudley. "Reflecting health: smart mirrors for personalized medicine". In: *NPJ digital medicine* 1.1 (2018), pp. 1–7.

[147] Arden Rowell. "Once and Future Nudges". In: *Missouri Law Review* 82 (2017), pp. 1–18.

[148] Guglielmo Briscese and Cameron Tan. *Applying Behavioural Insights to Labour Markets.* Report. The Behavioural Insights Team, 2018.

[149] Alexander J Stewart et al. "Information gerrymandering and undemocratic decisions". In: *Nature* 573.7772 (2019), pp. 117–121.

[150] Martin Adam and Mario Pecorelli. "Recommendations in augmented reality applications – The effect of customer reviews and seller recommendations on purchase in¬tention and product selection". In: *Proceedings of ECIS'18.* Portsmouth, UK: AIS, 2018, pp. 1–19.

[151] Martin Adam, Jonas Toutaoui, Nicolas Pfeuffer, and Oliver Hinz. "Investment decisions with robo-advisors: the role of anthropomorphism and personalized anchors in recommendations". In: *Proceedings of ECIS'19.* Stockholm & Uppsala, Sweden: AIS, 2019, pp. 1–18.

[152] Martin Adam, Michael Wessel, and Alexander Benlian. “Of early birds and phantoms: how sold-out discounts impact entrepreneurial success in reward-based crowdfunding”. In: *Review of Managerial Science* 13.3 (2019), pp. 545–560.

[153] Simge Andı and Jesper Akesson. “Nudging Away False News: Evidence from a Social Norms Experiment”. In: *Digital Journalism* 9:1 (2020), pp. 106–125.

[154] Marzia Antenore, Giovanna Leone, Alessandro Panconesi, and Erisa Terolli. “Together we buy, alone I quit: some experimental studies of online persuaders”. In: *Proceedings of DTUC’18*. Paris, France: ACM, 2018, pp. 1–4.

[155] Sandra Bammert, Ulrich Matthias König, Maximilian Roeglinger, and Tabitha Wruck. “Exploring potentials of digital nudging for business processes”. In: *Business Process Management Journal* 26 (2020), pp. 1329–1347.

[156] Michelle Berger, Chiara Müller, and Niclas Nüske. “Digital Nudging in Online Grocery Stores–Towards Ecologically Sustainable Nutrition”. In: *Proceedings of ICIS’20*. Hyderabad, India: AIS, 2020, pp. 1–17.

[157] Ivan Boskovic-Pavkovic, Isabella Seeber, and Ronald Maier. “How Digital Nudges Affect Consideration Set Size and Perceived Cognitive Effort in Idea Convergence of Open Innovation Contests”. In: *Proceedings of HICSS’19*. Maui, HI, USA: ScholarSpace, 2019, pp. 4415–4424.

[158] Yan Chen, Shuyuan Deng, Dong-Heon Kwak, Ahmed Elnoshokaty, and Jiao Wu. “A multi-appeal model of persuasion for online petition success: A linguistic cue-based approach”. In: *Journal of the Association for Information Systems* 20.2 (2019), pp. 105–131.

[159] Francis Joseph Costello, Jin Ho Yun, and Kun Chang Lee. “A NeuroIS Investigation of the Effects of a Digital Dark Nudge”. In: *NeuroIS Retreat*. Cham, Switzerland: Springer Nature, 2020, pp. 64–70.

[160] Nicolás E Dìaz Ferreyra, Tobias Kroll, Esma Aïmeur, Stefan Stieglitz, and Maritta Heisel. “Preventative Nudges: Introducing Risk Cues for Supporting Online Self-Disclosure Decisions”. In: *Information(Switzerland)* 11.8 (2020), p. 399.

[161] Marko Gasparic and Francesco Ricci. “IDE Interaction Support With Command Recommender Systems”. In: *IEEE Access* 8 (2020), pp. 19256–19270.

[162] Cristina Gena, Pierluigi Grillo, Antonio Lieto, Claudio Mattutino, and Fabiana Vernero. “When Personalization Is Not an Option: An In-The-Wild Study on Persuasive News Recommendation”. In: *Information (Switzerland)* 10.10 (2019), p. 300.

[163] Christopher Henkel, Anna-Raissa Seidler, Johann Kranz, and Marina Fiedler. “How to nudge pro-environmental behaviour: An experimental study”. In: *Proceedings of ECIS'19*. Stockholm and Uppsala, Sweden: AIS, 2019, p.

[164] Ni Huang, Peiyu Chen, Yili Hong, and Shinyi Wu. “Digital nudging for online social sharing: Evidence from a randomized field experiment”. In: *Proceedings of HICSS'18*. Waikoloa Village, HI, USA: AIS, 2018, pp. 1483–1491.

[165] Julian Huber, Dominik Jung, Elisabeth Schaule, and Christof Weinhardt. “Goal framing in smart charging-increasing BEV users' charging flexibility with digital nudges”. In: *Proceedings of ECIS'19*. Stockholm and Uppsala, Sweden: AIS, 2019, pp. 1–16.

[166] Shipi Kankane, Carlina DiRusso, and Christen Buckley. “Can We Nudge Users Toward Better Password Management? An Initial Study”. In: *Proceedings of CHI'18 Extended Abstracts*. Montreal, QC, Canada: ACM, 2018, pp. 1–6.

[167] Indrani Karmakar and Jane Webster. “Personalization of an Environmental Message: Developing a Measure”. In: *Proceedings of ICIS'18*. San Francisco, CA, USA: AIS, 2018, pp. 1–9.

[168] Katarzyna Katner and Radu Jianu. “The Effectiveness of Nudging in Commercial Settings and Impact on User Trust”. In: *Proceedings of CHI'19 Extended Abstracts*. Glasgow, Scotland, UK: ACM, 2019, pp. 1–6.

[169] Antino Kim, Patricia L Moravec, and Alan R Dennis. “Combating fake news on social media with source ratings: the effects of user and expert reputation ratings”. In: *Journal of Management Information Systems* 36.3 (2019), pp. 931–968.

[170] Tobias Kroll and Stefan Stieglitz. “Digital nudging and privacy: improving decisions about self-disclosure in social networks”. In: *Behaviour & Information Technology* 40:1 (2019), pp. 1–19.

[171] Tim-Benjamin Lembcke, Nils Engelbrecht, Mathias Willnat, and Sascha Lichtenberg. "Behavioral Design in Online Supermarkets: How Virtual Shopping Cart Functions Impact Sustainable Consumption". In: *Proceedings of AMCIS'20*. Salt Lake City, UT, USA: AIS, 2020, pp. 1–10.

[172] Yijing Li et al. "Exploring the Nudging and Counter-Nudging Effects of Campaign Updates in Crowdfunding". In: *Proceedings of PACIS'18*. Yokohama, Japan: AIS, 2018, pp. 1–8.

[173] Shouwang Lu, Gong Gordon Chen, and Kanliang Wang. "Overt or covert? Effect of different digital nudging on consumers' customization choices". In: *Nankai Business Review International* 12.1 (2020), pp. 56–74.

[174] Thomas Mejtoft, Charlotte Ristiniemi, Ulrik Söderström, and Eva Mårell-Olsson. "User experience design and digital nudging in a decision making process". In: *Proceedings of Bled eConference'19*. University of Maribor Press. Bled, Slovenia: AIS, 2019, pp. 427–442.

[175] Christian Meske, Ireti Amojo, and Peter Mohr. "Digital Nudging to Increase Usage of Charity Features on E-Commerce Platforms". In: *Proceedings of International Conference on Wirtschaftsinformatik*. Potsdam, Germany: AIS, 2020, pp. 1203–1218.

[176] Nuno Mota, Abhijnan Chakraborty, Asia J Biega, Krishna P Gummadi, and Hoda Heidari. "On the Desiderata for Online Altruism: Nudging for Equitable Donations". In: *Proceedings of the ACM on Human-Computer Interaction* 4.CSCW2 (2020), pp. 1–21.

[177] Gabriele Piccoli, Marcin Łukasz Bartosiak, Biagio Palese, and Joaquin Rodriguez. "Designing scalability in required in-class introductory college courses". In: *Information & Management* 57.8 (2020), p. 103263.

[178] Martin Poniatowski and Jürgen Neumann. "Getting Personal with Review Systems–Analyzing the Influence of Personality Traits on the Relationship between Review Templates and Reviewing Behavior". In: *Proceedings of ECIS'20 Research-in-Progress Papers*. Marrakech, Morocco: AIS, 2020, pp. 1–11.

[179] Aditya Kumar Purohit, Louis Barclay, and Adrian Holzer. "Designing for Digital Detox: Making Social Media Less Addictive with Digital Nudges".

In: *Proceedings of CHI'20 Extended Abstracts*. Honolulu, HI, USA: ACM, 2020, pp. 1–9.

[180] Joaquin Rodriguez, Gabriele Piccoli, and Marcin Bartosiak. "Nudging the classroom: Designing a socio-technical artifact to reduce academic procrastination". In: *Proceedings of HICSS'19*. Maui, HI, USA: ScholarSpace, 2019, pp. 4405–4414.

[181] Renata Santiago Walser, Isabella Seeber, and Ronald Maier. "Designing a Digital Nudge for Convergence: The Role of Decomposition of Information Load for Decision Making and Choice Accuracy". In: *AIS Transactions on Human-Computer Interaction* 11.3 (2019), pp. 179–207.

[182] David Schneider, Sebastian Lins, Tillmann Grupp, Alexander Benlian, and Ali Sunyaev. "Nudging users into online verification: the case of carsharing platforms". In: *Proceedings of ICIS'17*. Seoul, South Korea: AIS, 2017, pp. 1–20.

[183] Charlotte Schöning, Christian Matt, and Thomas Hess. "Personalised nudging for more data disclosure? On the adaption of data usage policies format to cognitive styles". In: *Proceedings of HICSS'19*. Maui, HI, USA: ScholarSpace, 2019, pp. 4395–4404.

[184] Tim Schrills, Mourad Zoubir, Jacob Stahl, Katharina Drozniak, and Thomas Franke. "Good Boy Here or Bad Boy Far Away?" In: *Proceedings of International Conference on Human-Computer Interaction*. Copenhagen, Denmark: Springer, 2020, pp. 398–411.

[185] Carola Stryja, Gerhard Satzger, and Verena Dorner. "A decision support system design to overcome resistance towards sustainable innovations". In: *Proceedings of ECIS'17 Research-in-Progress Papers*. Guimarães, Portugal: AIS, 2017, pp. 2885–2895.

[186] Carola Stryja and Gerhard Satzger. "Digital nudging to overcome cognitive resistance in innovation adoption decisions". In: *The Service Industries Journal* 39.15–16 (2019), pp. 1123–1139.

[187] Nadine Székely, Markus Weinmann, and Jan vom Brocke. "Nudging People To Pay Co2 Offsets–The Effect of Anchors in Flight Booking Processes". In: *Proceedings of ECIS'16 Research-in-Progress Papers*. Istanbul, Turkey: AIS, 2016, pp. 1–10.

[188] Philipp Terres, Johannes Klumpe, Dominik Jung, and Oliver Koch. "Digital Nudges for User Onboarding: Turning Visitors into Users". In: *Proceedings of ECIS'19*. Stockholm and Uppsala, Sweden: AIS, 2019, pp. 1–16.

[189] Arnold Wibmer, Frederik M Wiedmann, Isabella Seeber, and Ronald Maier. "Why less is more: An eye tracking study on idea presentation and attribute attendance in idea selection". In: *Proceedings of ECIS'19*. Stockholm and Uppsala, Sweden: AIS, 2019, pp. 1–14.

3. Nudging for Online Privacy Awareness and Informed Consent

[190] Jonathan A Obar and Anne Oeldorf-Hirsch. "The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services". In: *Information, Communication & Society* 23.1 (2020), pp. 128–147.

[191] H Jeff Smith, Tamara Dinev, and Heng Xu. "Information privacy research: an interdisciplinary review". In: *MIS quarterly* 35.4 (2011), pp. 989–1016.

[192] Tamara Dinev, Allen R McConnell, and H Jeff Smith. "Research commentary—informing privacy research through information systems, psychology, and behavioral economics: thinking outside the "APCO" box". In: *Information Systems Research* 26.4 (2015), pp. 639–655.

[193] S. Furnell and A. Phippen. "Online privacy: a matter of policy?" In: *Computer Fraud & Security* 2012.8 (2012), pp. 12–18. ISSN: 1361-3723.

[194] Aleecia M McDonald and Lorrie Faith Cranor. "The cost of reading privacy policies". In: *ISJLP* 4 (2008), p. 543.

[195] T. Ermakova, A. Baumann, B. Fabian, and H. Krasnova. "Privacy Policies and Users' Trust: Does Readability Matter?" In: *Twentieth americas conference on information systems*. Savanna, 2014, pp. 1–12.

[196] Spyros Kokolakis. "Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon". In: *Computers & security* 64 (2017), pp. 122–134.

[197] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. "Privacy and human behavior in the age of information". In: *Science* 347.6221 (2015), pp. 509–514.

[198] A. Acquisti. "Nudging privacy: The behavioral economics of personal information". In: *IEEE security & privacy* 7.6 (2009), pp. 82–85.

[199] J. Y. Tsai, S. Egelman, L. F. Cranor, and A. Acquisti. "The effect of online privacy information on purchasing behavior: An experimental study". In: *Information Systems Research* 22.2 (2011), pp. 254–268.

[200] Hichang Cho, Jae-Shin Lee, and Siyoung Chung. "Optimistic bias about online privacy risks: Testing the moderating effects of perceived controllability and prior experience". In: *Computers in Human Behavior* 26.5 (2010), pp. 987–995.

[201] Yang Wang et al. "" I regretted the minute I pressed share" a qualitative study of regrets on Facebook". In: *Proceedings of the seventh symposium on usable privacy and security*. 2011, pp. 1–16.

[202] Alessandro Acquisti et al. "Nudges for privacy and security: Understanding and assisting users' choices online". In: *ACM Computing Surveys (CSUR)* 50.3 (2017), pp. 1–41.

[203] Helia Marreiros, Mirco Tonin, Michael Vlassopoulos, and MC Schraefel. ""Now that you mention it": A survey experiment on information, inattention and online privacy". In: *Journal of Economic Behavior & Organization* 140 (2017), pp. 1–17.

[204] Chris Jay Hoofnagle and Jennifer M Urban. "Alan Westin's privacy homo economicus". In: *Wake Forest L. Rev.* 49 (2014), p. 261.

[205] J. P. Carrascal, C. Riederer, V. Erramilli, M. Cherubini, and R. de Oliveira. "Your browsing behavior for a big mac: Economics of personal information online". In: *Proceedings of the 22nd international conference on World Wide Web*. ACM, 2013, pp. 189–200.

[206] Richard H Thaler and Cass R Sunstein. *Nudge: Improving decisions about health, wealth, and happiness*. Penguin, 2009.

[207] Yang Wang et al. "Privacy nudges for social media: an exploratory Facebook study". In: *Proceedings of the 22nd International Conference on World Wide Web*. 2013, pp. 763–770.

[208] Marian Harbach, Markus Hettig, Susanne Weber, and Matthew Smith. "Using personal examples to improve risk communication for security & privacy decisions". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2014, pp. 2647–2656.

[209] The Internet Archive. *Connect with friends and the world around you on Facebook*. Website, Accessed on: 02/12/2023. Available at: `https://web.archive.org/web/20191101003352/https://www.facebook.com/`. 2019.

[210] The Internet Archive. *Make the Most of your Professional Life*. Website, Accessed on: 02/12/2023. Available at: `https://web.archive.org/web/20191111185524/https://www.linkedin.com/start/join?trk=guest_homepage-basic_nav-header-join`. 2019.

[211] The Internet Archive. *Create your Account*. Website, Accessed on: 21/11/2023. Available at: `https://web.archive.org/web/20191101013136/https://twitter.com/i/flow/signup`. 2019.

[212] Ken Peffers, Tuure Tuunanen, Marcus A Rothenberger, and Samir Chatterjee. "A design science research methodology for information systems research". In: *Journal of management information systems* 24.3 (2007), pp. 45–77.

[213] R. G. Kunkel. "Recent developments in shrinkwrap, clickwrap and browsewrap licenses in the United States". In: *Murdoch University Electronic Journal of Law* 9.3 (2002).

[214] Gloria E Phillips-Wren, Lakshmi S Iyer, Uday R Kulkarni, and Thilini Ariyachandra. "Business Analytics in the Context of Big Data: A Roadmap for Research." In: *CAIS* 37 (2015), p. 23.

[215] Robert A Hillman and Jeffrey J Rachlinski. "Standard-form contracting in the electronic age". In: *NYUL Rev.* 77 (2002), p. 429.

[216] Jonathan A Obar and Anne Oeldorf-Hirsch. "Clickwrap impact: Quick-join options and ignoring privacy and terms of service policies of social networking services". In: *Proceedings of the 8th International Conference on Social Media & Society*. 2017, pp. 1–5.

[217] Partha Dasgupta and Eric Maskin. "Uncertainty and hyperbolic discounting". In: *American Economic Review* 95.4 (2005), pp. 1290–1299.

[218] Ted O'Donoghue and Matthew Rabin. "The economics of immediate gratification". In: *Journal of Behavioral Decision Making* 13.2 (2000), pp. 233–250.

[219] Alessandro Acquisti and Jens Grossklags. “Losses, gains, and hyperbolic discounting: An experimental approach to information security attitudes and behavior”. In: *2nd Annual Workshop on Economics and Information Security-WEIS*. Vol. 3. 2003, pp. 1–27.

[220] Alessandro Acquisti and Jens Grossklags. “Privacy and rationality in individual decision making”. In: *IEEE security & privacy* 3.1 (2005), pp. 26–33.

[221] A. K Purohit, L Barclay, and A. Holzer. “Designing for Digital Detox: Making Social Media Less Addictive with Digital Nudges”. In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI EA ’20. Honolulu, HI, USA: ACM, 2020, pp. 1–9.

[222] Brian J Fogg. “Persuasive computers: perspectives and research directions”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co. 1998, pp. 225–232.

[223] Brian J Fogg. “A behavior model for persuasive design”. In: *Proceedings of the 4th international Conference on Persuasive Technology*. ACM, 2009, pp. 1–7.

[224] Eric J Johnson and Daniel Goldstein. “Do defaults save lives?” In: *Science* 302.5649 (2003), p. 1338.

[225] Joel B Cohen, Eduardo B Andrade, et al. “The ADF Framework: A Parsimonious Model for Developing Successful Behavior Change Interventions”. In: *Journal of Marketing Behavior* 3.2 (2018), pp. 81–119.

[226] Alain Forget, Sonia Chiasson, Paul C Van Oorschot, and Robert Biddle. “Improving text passwords through persuasion”. In: *Proceedings of the 4th symposium on Usable privacy and security*. ACM. 2008, pp. 1–12.

[227] Eric P Bettinger, Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. “The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment”. In: *The Quarterly Journal of Economics* 127.3 (2012), pp. 1205–1242.

[228] Suzanna E Forwood, Amy L Ahern, Theresa M Marteau, and Susan A Jebb. “Offering within-category food swaps to reduce energy density of food purchases: a study using an experimental online supermarket”. In: *International Journal of Behavioral Nutrition and Physical Activity* 12.1 (2015), p. 85.

[229] Alexander Simons, Markus Weinmann, Matthias Tietz, and Jan vom Brocke. "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*. 2017.

[230] Itamar Simonson. "Choice based on reasons: The case of attraction and compromise effects". In: *Journal of consumer research* 16.2 (1989), pp. 158–174.

[231] Brian J Fogg. *Fogg Behavior Model*. Website. Available at: `https://www.behaviormodel.org/`. Sept. 2019.

[232] Gabriele Paolacci, Jesse Chandler, and Panagiotis G Ipeirotis. "Running experiments on amazon mechanical turk". In: *Judgment and Decision making* 5.5 (2010), pp. 411–419.

[233] Jan H Schumann, Florian von Wangenheim, and Nicole Groene. "Targeted online advertising: Using reciprocity appeals to increase acceptance among users of free web services". In: *Journal of Marketing* 78.1 (2014), pp. 59–75.

[234] Daniel J Benjamin et al. "Redefine statistical significance". In: *Nature Human Behaviour* 2.1 (2018), p. 6.

[235] Kristen L Walker. "Surrendering information through the looking glass: Transparency, trust, and protection". In: *Journal of Public Policy & Marketing* 35.1 (2016), pp. 144–158.

[236] Min K Lee, Sara Kiesler, and Jodi Forlizzi. "Mining behavioral economics to design persuasive technology for healthy choices". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2011, pp. 325–334.

[237] Christoph Buck. "Stop disclosing personal data about your future self". In: *23th Americas Conference on Information Systems (AMCIS)*. 2017, pp. 1–10.

[238] David J Hauser and Norbert Schwarz. "Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants". In: *Behavior research methods* 48.1 (2016), pp. 400–407.

4. Reducing a Newsfeed for Social Media Detox

[239] Global Web Index. *Flagship Report 2019*. Report. Global Web Index, 2019.

[240] Antti Oulasvirta, Tye Rattenbury, Lingyi Ma, and Eeva Raita. “Habits make smartphone use more pervasive”. In: *Personal and Ubiquitous Computing* 16.1 (2012), pp. 105–114.

[241] Joseph B. Bayer, Ian A. Anderson, and Robert S. Tokunaga. “Building and breaking social media habits”. In: *Current Opinion in Psychology* 45 (2022), p. 101303. ISSN: 2352-250X.

[242] Klodiana Lanaj, Russell E. Johnson, and Christopher M. Barnes. “Beginning the workday yet already depleted? Consequences of late-night smartphone use and sleep”. In: *Organizational Behavior and Human Decision Processes* 124.1 (2014), pp. 11–23.

[243] Joshua C Watson, Elizabeth A Prosek, and Amanda L Giordano. “Distress among adolescents: An exploration of mattering, social media addiction, and school connectedness”. In: *Journal of Psychoeducational Assessment* 40.1 (2022), pp. 95–107.

[244] Yu-Kang Lee, Chun-Tuan Chang, You Lin, and Zhao-Hong Cheng. “The dark side of smartphone usage: Psychological traits, compulsive behavior and technostress”. In: *Computers in Human Behavior* 31 (2014), pp. 373–383.

[245] Chuang Wang and Matthew KO Lee. “Why we cannot resist our smartphones: investigating compulsive use of mobile SNS from a Stimulus-Response-Reinforcement perspective”. In: *Journal of the Association for Information Systems* 21.1 (2020), p. 4.

[246] Simon Kloker. “Non-addictive information systems”. In: *Information Systems Frontiers* 22 (2020), pp. 549–562.

[247] Santiago Giraldo-Luque, Pedro Nicolás Aldana Afanador, and Cristina Fernández-Rovira. “The Struggle for Human Attention: Between the Abuse of Social Media and Digital Wellbeing”. In: *Healthcare* 8.4 (2020), p. 497.

[248] Vikram R Bhargava and Manuel Velasquez. "Excerpt from Ethics of the Attention Economy: The Problem of Social Media Addiction". In: *Ethics of Data and Analytics*. New York, NY, USA: Auerbach Publications, 2022, pp. 391–402.

[249] Alberto Monge R. and Luigi De Russis. "Towards Understanding the Dark Patterns That Steal Our Attention". In: *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. CHI EA '22. New York, NY, USA: ACM, 2022, pp. 1–7.

[250] Philippe Verduyn et al. "Passive Facebook usage undermines affective well-being: Experimental and longitudinal evidence." In: *Journal of Experimental Psychology: General* 144.2 (2015), p. 480.

[251] Shirley Gregor, David Jones, et al. "The anatomy of a design theory". In: *Journal of the Association for Information Systems* 8.5 (2007), pp. 312–335.

[252] Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. "Design Science in Information Systems Research". In: *MIS Quarterly* 28.1 (2004), pp. 75–105. ISSN: 02767783.

[253] Alan R Hevner. "A three cycle view of design science research". In: *Scandinavian journal of information systems* 19.2 (2007), pp. 87–92.

[254] Jan Holmström, Mikko Ketokivi, and Ari-Pekka Hameri. "Bridging practice and theory: A design science approach". In: *Decision sciences* 40.1 (2009), pp. 65–87.

[255] Simone Lanette, Phoebe K Chua, Gillian Hayes, and Melissa Mazmanian. "How much is' too much'? The role of a smartphone addiction narrative in individuals' experience of use". In: *Proceedings of the ACM on Human-Computer Interaction* 2.CSCW (2018), pp. 1–22.

[256] Jonathan A. Tran, Katie S. Yang, Katie Davis, and Alexis Hiniker. "Modeling the Engagement-Disengagement Cycle of Compulsive Phone Use". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: ACM, 2019, pp. 1–14.

[257] Amanda Baughan et al. ""I Don't Even Remember What I Read": How Design Influences Dissociation on Social Media". In: *CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, 2022, pp. 1–13.

[258] Kai Lukoff, Cissy Yu, Julie Kientz, and Alexis Hiniker. "What makes smartphone use meaningful or meaningless?" In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2.1 (2018), pp. 1–26.

[259] Kai Lukoff et al. "How the design of youtube influences user sense of agency". In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: ACM, 2021, pp. 1–17.

[260] Jane E Klobas, Tanya J McGill, Sedigheh Moghavvemi, and Tanousha Paramanathan. "Compulsive YouTube usage: A comparison of use motivation and personality effects". In: *Computers in Human Behavior* 87 (2018), pp. 129–139.

[261] Mark D. Griffiths, Daria J. Kuss, and Zsolt Demetrovics. "Chapter 6 - Social Networking Addiction: An Overview of Preliminary Findings". In: *Behavioral Addictions*. Ed. by Kenneth Paul Rosenberg and Laura Curtiss Feder. San Diego: Academic Press, 2014, pp. 119–141. ISBN: 978-0-12-407724-9.

[262] Emma Fontes-Perryman and Roy Spina. "Fear of missing out and compulsive social media use as mediators between OCD symptoms and social media fatigue." In: *Psychology of Popular Media* 11.2 (2022), p. 173.

[263] Vladlena Benson, Chris Hand, and Richard Hartshorne. "How compulsive use of social media affects performance: insights from the UK by purpose of use". In: *Behaviour & Information Technology* 38.6 (2019), pp. 549–563.

[264] Christian Montag, Bernd Lachmann, Marc Herrlich, and Katharina Zweig. "Addictive features of social media / messenger platforms and freemium games against the background of psychological and economic theories". In: *International journal of environmental research and public health* 16.14 (2019), p. 2612.

[265] Siddhartha Datta, Konrad Kollnig, and Nigel Shadbolt. "Mind-proofing Your Phone: Navigating the Digital Minefield with GreaseTerminator". In: *27th International Conference on Intelligent User Interfaces*. New York, NY, USA: ACM, 2022, pp. 523–536.

[266] Ulrik Lyngs et al. "'I Just Want to Hack Myself to Not Get Distracted' Evaluating Design Interventions for Self-Control on Facebook". In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: ACM, 2020, pp. 1–15.

[267] Young-Ho Kim et al. "TimeAware: Leveraging Framing Effects to Enhance Personal Productivity". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. San Jose, California, USA: ACM, 2016, pp. 272–283.

[268] Sarawut Thongsrikum. "The Art of Screen Time: How Your Family Can Balance Digital Media and Real Life". In: *Asian Journal of Arts and Culture* 21.2 (2021), pp. 71–73.

[269] Ulrik Lyngs et al. "The Goldilocks level of support: Using user reviews, ratings, and installation numbers to investigate digital self-control tools". In: *International journal of human-computer studies* 166 (2022), p. 102869.

[270] Ulrik Lyngs et al. "Self-control in cyberspace: Applying dual systems theory to a review of digital self-control tools". In: *Proceedings of the 2019 CHI conference on human factors in computing systems*. New York, NY, USA: ACM, 2019, pp. 1–18.

[271] Stephanie Nguyen and Jasmine McNealy. *I, Obscura — Illuminating deceptive design patterns in the wild*. Available at: `https://issuu.com/stanforddcsl/docs/dcsl_darkpatternszine_2021`. 2021.

[272] Aditya K. Purohit and Adrian Holzer. "Unhooked by Design: Scrolling Mindfully on Social Media by Automating Digital Nudges". In: *Proceeding of the 27th Americas Conference on Information Systems, AMCIS 2021*. Virtual Conference: AIS, 2021, pp. 1–10.

[273] Wendy Wood and David T Neal. "The habitual consumer". In: *Journal of Consumer Psychology* 19.4 (2009), pp. 579–592.

[274] Unna N Danner, Henk Aarts, and Nanne K De Vries. "Habit vs. intention in the prediction of future behaviour: The role of frequency, context stability and mental accessibility of past behaviour". In: *British Journal of Social Psychology* 47.2 (2008), pp. 245–265.

[275] Agata Błachnio and Aneta Przepiórka. "Facebook intrusion, fear of missing out, narcissism, and life satisfaction: A cross-sectional study". In: *Psychiatry Research* 259 (2018), pp. 514–519. ISSN: 0165-1781.

[276] Mingrui Ray Zhang, Kai Lukoff, Raveena Rao, Amanda Baughan, and Alexis Hiniker. "Monitoring Screen Time or Redesigning It? Two Approaches to Supporting Intentional Social Media Use". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. New Orleans, LA, USA: ACM, 2022, pp. 1–19.

[277] Edward L. Deci and Richard M. Ryan. "Self-Determination Theory". In: *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*. Ed. by James D. Wright. Second Edition. Oxford: Elsevier, 2015, pp. 486–491.

[278] Hannah Limerick, David Coyle, and James W Moore. "The experience of agency in human-computer interactions: a review". In: *Frontiers in human neuroscience* 8 (2014), p. 643.

[279] Meta. *Meta Reports Fourth Quarter and Full Year 2021 Results*. Website, Accessed on: 05/07/2022. Available at: `https://investor.fb.com/investor-news/press-release-details/2022/Meta-Reports-Fourth-Quarter-and-Full-Year-2021-Results/default.aspx`. Feb. 2022.

[280] Meta. *How does Facebook decide which ads to show me?* Website, Accessed on: 05/07/2022. Available at: `https://www.facebook.com/help/562973647153813`. 2022.

[281] Alberto Monge Roffarello and Luigi De Russis. "Coping with digital wellbeing in a multi-device world". In: *Proceedings of the CHI conference on human factors in computing cystems*. CHI '21. New York, NY, USA: ACM, 2021, pp. 1–14.

[282] Facebook. *Tools*. Retrieved January 30, 2023 from: `https://www.facebook.com/safety/tools/safety`. 2023.

[283] Aaron Smith. *What people like and dislike about Facebook*. Pew Research Center. Feb. 2014. URL: `http://pewrsr.ch/1dm5NmJ`.

[284] Aditya Kumar Purohit, Torben Jan Barev, Sofia Schöbel, Andreas Janson, and Adrian Holzer. "Designing for Digital Wellbeing on a Smartphone: Co-creation of Digital Nudges to Mitigate Instagram Overuse". In: *Proceedings of the 56th Hawaii International Conference on System Sciences*. HICSS '56. Hawai'i, USA: University of Hawai'i at Mānoa, 2023, pp. 4087–4096.

[285] Highcharts. *Highcharts Javascript Charting Library*. Website, Accessed on: 22/09/2021. Available at: `https://www.highcharts.com/products/highcharts/`. 2021.

[286] Global Web Index. *Social - GWI's flagship report on the latest trends in social media*. Report. Global Web Index, 2021.

[287] Micah Allen, Davide Poggiali, Kirstie Whitaker, Tom Rhys Marshall, and Rogier A Kievit. "Raincloud plots: a multi-platform tool for robust data visualization [version 2; peer review: 2 approved]". In: *Wellcome open research* 4.63 (2021), pp. 1–40.

[288] Micah Allen, Davide Poggiali, Kirstie Whitaker, T Rhys Marshall, and Rogier A Kievit. *RainCloudPlots tutorials and codebase (Version 1.1). Zenodo*. Retrieved November 14, 2022 from: `https://zenodo.org/record/3368186#.Y3yQmnbMIuU`. 2019.

[289] Mikel Aickin. "Dealing with change: using the conditional change model for clinical research". In: *The Permanente Journal* 13.2 (2009), p. 80.

[290] Kevin A Hallgren. "Computing inter-rater reliability for observational data: an overview and tutorial". In: *Tutorials in quantitative methods for psychology* 8.1 (2012), p. 23.

[291] Nahathai Wongpakaran, Tinakon Wongpakaran, Danny Wedding, and Kilem L Gwet. "A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples". In: *BMC medical research methodology* 13.1 (2013), pp. 1–7.

[292] J. Saldaña. *The Coding Manual for Qualitative Researchers*. Core textbook. London: SAGE, 2021. ISBN: 9781529731743. URL: `https://books.google.nl/books?id=iQK1zQEACAAJ`.

[293] Virginia Braun and Victoria Clarke. "Using thematic analysis in psychology". In: *Qualitative Research in Psychology* 3.2 (2006), pp. 77–101.

[294] Jennifer Fereday and Eimear Muir-Cochrane. "Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development". In: *International journal of qualitative methods* 5.1 (2006), pp. 80–92.

[295] Elisa Wegmann, Ursula Oberst, Benjamin Stodt, and Matthias Brand. "Online-specific fear of missing out and Internet-use expectancies contribute to symptoms of Internet-communication disorder". In: *Addictive Behaviors Reports* 5 (2017), pp. 33–42.

[296] Teun Siebers, Ine Beyens, J Loes Pouwels, and Patti M Valkenburg. "Social media and distraction: An experience sampling study among adolescents". In: *Media Psychology* 25.3 (2022), pp. 343–366.

[297] Caio A. Lage, De Wet Wolmarans, and Daniel C. Mograbi. "An evolutionary view of self-awareness". In: *Behavioural Processes* 194 (2022), p. 104543.

[298] Joseph B. Bayer, Scott W. Campbell, and Rich Ling. "Connection Cues: Activating the Norms and Habits of Social Connectedness". In: *Communication Theory* 26.2 (Nov. 2015), pp. 128–149.

[299] Jin P Gerlach and Ronald T Cenfetelli. "Constant Checking Is Not Addiction: A Grounded Theory of IT-Mediated State-Tracking". In: *MIS Quarterly* 44 (4 2020), pp. 1705–1732.

[300] C. Crumlish and E. Malone. *Designing Social Interfaces: Principles, Patterns, and Practices for Improving the User Experience*. Animal Guide. Canada: O'Reilly Media, 2009. ISBN: 9781449379223.

[301] James W Moore. "What is the sense of agency and why does it matter?" In: *Frontiers in psychology* 7 (2016), p. 1272.

[302] Carolyn R Plateau, Trent A Petrie, and Anthony Papathomas. "Learning to eat again: Intuitive eating practices among retired female collegiate athletes". In: *Eating disorders* 25.1 (2017), pp. 92–98.

[303] Minsam Ko, Seungwoo Choi, Koji Yatani, and Uichin Lee. "Lock n' LoL: Group-Based Limiting Assistance App to Mitigate Smartphone Distractions in Group Activities". In: *Proceedings of the CHI conference on human factors in computing systems*. CHI '16. San Jose, California, USA: ACM, 2016, pp. 998–1010.

[304] Jennifer Xu, Jane Fedorowicz, and Christine B Williams. "Effects of symbol sets and needs gratifications on audience engagement: Contextualizing police social media communication". In: *Journal of the Association for Information Systems* 20.5 (2019), pp. 308–327.

[305] Jaejeung Kim, Hayoung Jung, Minsam Ko, and Uichin Lee. "GoalKeeper: Exploring Interaction Lockout Mechanisms for Regulating Smartphone Use". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3.1 (2019), pp. 1–29.

[306] Alex Heath. *Facebook is changing its algorithm to take on TikTok, leaked memo reveals*. Retrieved February 15, 2023 from: https://www.theverge.com/2022/6/15/23168887/facebook-discovery-engine-redesign-tiktok. 2022.

[307] Cal Newport. *TikTok and the Fall of the Social-Media Giants*. Available at: https://www.newyorker.com/culture/cultural-comment/tiktok-and-the-fall-of-the-social-media-giants. 2022.

[308] MR Leiser and M Caruana. "Dark Patterns: Light to be found in Europe's Consumer Protection Regime". In: *Journal of European Consumer and Market Law* 10.6 (2021), pp. 237–251.

[309] David Coyle, James Moore, Per O. Kristensson, Paul Fletcher, and Alan Blackwell. "I Did That! Measuring Users' Experience of Agency in Their Own Actions". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '12. Austin, Texas, USA: ACM, 2012, pp. 2025–2034.

[310] Ben Shneiderman. "Shneiderman's eight golden rules of interface design". In: *Retrieved july* 25 (2005), p. 2009.

5. Stacking Nudges for Online Charitable Giving

[311] Philanthropy Roundtable. *Statistics on U.S. Generosity*. Website, Accessed on: 04/10/2023. Available at: `https://www.philanthropyroundtable.org/almanac/statistics-on-u-s-generosity/`. 2023.

[312] Indiana University (IUPUI): Lilly Family School of Philanthropy. *USA 2023: The Annual Report on Philanthropy for the Year 2022*. Tech. rep. Giving USA Foundation, 2023.

[313] Indiana University (IUPUI): Lilly Family School of Philanthropy. *USA 2022: The Annual Report on Philanthropy for the Year 2021*. Tech. rep. Giving USA Foundation, 2022.

[314] Blackbaud Institute. *Online Giving Trends*. Website, Accessed on: 02/11/2023. Available at: `https://institute.blackbaud.com/charitable-giving-report/online-giving-trends/`. 2022.

[315] Indiana University (IUPUI): Lilly Family School of Philanthropy. *Digital For Good: A Global Study on Emerging Ways of Giving*. Tech. rep. 2023.

[316] Nonprofit Tech for Good. *Global NGO Technology Report 2019*. Tech. rep. 2019.

[317] Rebecca C Ruehle, Bart Engelen, and Alfred Archer. "Nudging charitable giving: What (if anything) is wrong with it?" In: *Nonprofit and Voluntary Sector Quarterly* 50.2 (2021), pp. 353–371.

[318] William Hagman, David Andersson, Daniel Västfjäll, and Gustav Tinghög. "Public views on policies involving nudges". In: *Review of philosophy and psychology* 6 (2015), pp. 439–453.

[319] Alexis Shore and James J Cummings. "An Exploratory Analysis of Interface Features Influencing Mobile Location Data Disclosure". In: *International Journal of Human–Computer Interaction* (2022), pp. 1–16.

[320] Steffen Altmann, Armin Falk, Paul Heidhues, Rajshri Jayaraman, and Marrit Teirlinck. “Defaults and donations: Evidence from a field experiment”. In: *Review of Economics and Statistics* 101.5 (2019), pp. 808–826.

[321] Erik Gråd, Arvid Erlandsson, and Gustav Tinghög. “Do nudges crowd out prosocial behavior?” In: *Behavioural Public Policy* (2021), pp. 1–14.

[322] Valerio Capraro, Glorianna Jagfeld, Rana Klein, Mathijs Mul, and Iris van de Pol. “Increasing altruistic and cooperative behaviour with simple moral nudges”. In: *Scientific reports* 9.1 (2019), p. 11880.

[323] Jim AC Everett, Lucius Caviola, Guy Kahane, Julian Savulescu, and Nadira S Faber. “Doing good by doing nothing? The role of social norms in explaining default effects in altruistic contexts”. In: *European Journal of Social Psychology* 45.2 (2015), pp. 230–241.

[324] Christophe Charlier, Gilles Guerassimoff, Ankinée Kirakosian, and Sandrine Selosse. “Under pressure! Nudging electricity consumption within firms. Feedback from a field experiment”. In: *The Energy Journal* 41.1 (2021), pp. 129–154.

[325] Leilei Qu, Ruojin Xiao, and Wenchang Shi. “Interactions of Framing and Timing in Nudging Online Game Security”. In: *Computers & Security* 124 (2023), p. 102962.

[326] Mette Trier Damgaard and Christina Gravert. “The hidden costs of nudging: Experimental evidence from reminders in fundraising”. In: *Journal of Public Economics* 157 (2018), pp. 15–26.

[327] Sünje Clausen, Julian Marx, Milad Mirbabaie, and Stefan Stieglitz. “From Dark Patterns to Digital Sludging–Mapping the Ethical Debate on Controversial Persuasive System Design”. In: *Proceedings of the International Conference on Information Systems (ICIS)*. AIS, 2022, pp. 1–9.

[328] Daniel Susser and Vincent Grimaldi. “Measuring automated influence: between empirical evidence and ethical values”. In: *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*. 2021, pp. 242–253.

[329] Richard L Baskerville, Mala Kaul, and Veda C Storey. “Genres of inquiry in design-science research”. In: *Mis Quarterly* 39.3 (2015), pp. 541–564.

[330] Milad Mirbabaie, Julian Marx, and Lukas Erle. "Digital Nudge Stacking and Backfiring: Understanding Sustainable E-Commerce Purchase Decisions". In: *Pacific Asia Journal of the Association for Information Systems* 15.3 (2023), pp. 65–104.

[331] Anastasia Vugts, Mariëtte Van Den Hoven, Emely De Vet, and Marcel Verweij. "How autonomy is understood in discussions on the ethics of nudging". In: *Behavioural Public Policy* 4.1 (2020), pp. 108–123.

[332] Cass R Sunstein. "Nudging and choice architecture: Ethical considerations". In: *Yale Journal on Regulation, Forthcoming* 809 (2015).

[333] T Martin Wilkinson. "Nudging and manipulation". In: *Political Studies* 61.2 (2013), pp. 341–355.

[334] Cass R Sunstein. "Fifty shades of manipulation". In: *Journal of Marketing Behavior* 1.3–4 (2016), pp. 213–244.

[335] Alexios Arvanitis, Konstantinos Kalliris, and Konstantinos Kaminiotis. "Are defaults supportive of autonomy? An examination of nudges under the lens of Self-Determination Theory". In: *The Social Science Journal* 59.3 (2022), pp. 394–404.

[336] Patrik Michaelsen, Lars-Olof Johansson, and Martin Hedesström. "Experiencing default nudges: autonomy, manipulation, and choice-satisfaction as judged by people themselves". In: *Behavioural Public Policy* (2021), pp. 1–22.

[337] Stephanie Mertens, Mario Herberz, Ulf JJ Hahnel, and Tobias Brosch. "The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains". In: *Proceedings of the National Academy of Sciences* 119.1 (2022), e2107346118.

[338] Jon M Jachimowicz, Shannon Duncan, Elke U Weber, and Eric J Johnson. "When and why defaults influence decisions: A meta-analysis of default effects". In: *Behavioural Public Policy* 3.2 (2019), pp. 159–186.

[339] William Samuelson and Richard Zeckhauser. "Status quo bias in decision making". In: *Journal of risk and uncertainty* 1 (1988), pp. 7–59.

[340] Daniel Kahneman, Jack L Knetsch, and Richard H Thaler. "Anomalies: The endowment effect, loss aversion, and status quo bias". In: *Journal of Economic perspectives* 5.1 (1991), pp. 193–206.

[341] Klaus Goffart, Michael Schermann, Christopher Kohl, Jörg Preißinger, and Helmut Krcmar. "Using the default option bias to influence decision making while driving". In: *International Journal of Human-Computer Interaction* 32.1 (2016), pp. 39–50.

[342] Claus Ghesla, Manuel Grieder, and Jan Schmitz. "Nudge for good? Choice defaults and spillover effects". In: *Frontiers in psychology* 10 (2019), p. 178.

[343] Homa S Zarghamee et al. "Nudging charitable giving: Three field experiments". In: *Journal of behavioral and experimental economics* 66 (2017), pp. 137–149.

[344] Clark Leonard Hull. *Principles of behavior: an introduction to behavior theory*. NY, USA: Appleton-Century Crofts, 1943.

[345] Anna L Cox, Sandy JJ Gould, Marta E Cecchinato, Ioanna Iacovides, and Ian Renfree. "Design frictions for mindful interactions: The case for microboundaries". In: *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems*. ACM New York, NY, USA, 2016, pp. 1389–1397.

[346] Ron Van Houten, Paul A Nau, and Michael Merrigan. "Reducing elevator energy use: A comparison of posted feedback and reduced elevator convenience". In: *Journal of Applied Behavior Analysis* 14.4 (1981), pp. 377–387.

[347] Jooyeon Kim, Behzad Tabibian, Alice Oh, Bernhard Schölkopf, and Manuel Gomez-Rodriguez. "Leveraging the crowd to detect and reduce the spread of fake news and misinformation". In: *Proceedings of the eleventh ACM international conference on web search and data mining*. ACM New York, NY, USA, 2018, pp. 324–332.

[348] Juanjuan Zhang. "The sound of silence: Observational learning in the US kidney market". In: *Marketing Science* 29.2 (2010), pp. 315–335.

[349] Hunt Allcott and Judd B Kessler. "The welfare effects of nudges: A case study of energy use social comparisons". In: *American Economic Journal: Applied Economics* 11.1 (2019), pp. 236–276.

[350] Magda Osman et al. "Learning from behavioural changes that fail". In: *Trends in Cognitive Sciences* 24.12 (2020), pp. 969–980.

[351] Vladimir Melnyk, François A Carrillat, and Valentyna Melnyk. "The influence of social norms on consumer behavior: A meta-analysis". In: *Journal of Marketing* 86.3 (2022), pp. 98–120.

[352] International Committee of the Red Cross (ICRC). *Donate to the ICRC*. Website, Accessed on: 06/11/2023. Available at: `https://www.icrc.org/en/donate`. 2023.

[353] Emily Geisen. *Improve data quality by using a commitment request instead of attention checks*. Website, Accessed on: 29/03/2023. Available at: `https://www.qualtrics.com/blog/attention-checks-and-data-quality/`. 2022.

[354] James Price Dillard and Lijiang Shen. "On the nature of reactance and its role in persuasive health communication". In: *Communication monographs* 72.2 (2005), pp. 144–168.

[355] Kim Witte. "Fear control and danger control: A test of the extended parallel process model (EPPM)". In: *Communications Monographs* 61.2 (1994), pp. 113–134.

[356] Jeff Sauro. "SUPR-Q: A comprehensive measure of the quality of the website user experience." In: *Journal of usability studies* 10.2 (2015), pp. 68–86.

[357] Deborah J Webb, Corliss L Green, and Thomas G Brashear. "Development and validation of scales to measure attitudes influencing monetary donations to charitable organizations". In: *Journal of the academy of marketing science* 28.2 (2000), pp. 299–309.

[358] William O Bearden, Richard G Netemeyer, and Kelly L Haws. *Handbook of marketing scales: Multi-item measures for marketing and consumer behavior research*. Sage Publications, Sage CA: Thousand Oaks, CA, 2011, p. 603.

[359] Maaike Roubroeks, Jaap Ham, and Cees Midden. "When artificial social agents try to persuade people: The role of social agency on the occurrence of psychological reactance". In: *International Journal of Social Robotics* 3 (2011), pp. 155–165.

[360] Nir Eyal. *Want to Design User Behavior? Pass the 'Regret Test' First*. Accessed on 02/06/2023 at: `https://www.nirandfar.com/regret-test/`. 2022.

[361] Scott Menard. *Applied logistic regression analysis*. Thousand Oaks, CA, USA: Sage, 2002.

[362] Eric Vittinghoff, David V Glidden, Stephen C Shiboski, and Charles E McCulloch. *Regression methods in biostatistics: linear, logistic, survival, and repeated measures models*. Spring Street, NY, USA: Springer, 2006.

[363] Richard E Boyatzis. *Transforming qualitative information: Thematic analysis and code development*. sage, 1998.

[364] J Richard Landis and Gary G Koch. "The measurement of observer agreement for categorical data". In: *biometrics* (1977), pp. 159–174.

[365] Indranil Goswami and Oleg Urminsky. "When should the ask be a nudge? The effect of default amounts on charitable donations". In: *Journal of Marketing Research* 53.5 (2016), pp. 829–846.

[366] Gary E Bolton, Eugen Dimant, and Ulrich Schmidt. *When a nudge backfires: Using observation with social and economic incentives to promote pro-social behavior*. Tech. rep. CeDEx Discussion Paper Series, 2019.

[367] Terry Connolly and Marcel Zeelenberg. "Regret in decision making". In: *Current directions in psychological science* 11.6 (2002), pp. 212–216.

[368] Peter Allmark. "Should research samples reflect the diversity of the population?" In: *Journal of medical ethics* 30.2 (2004), pp. 185–189.

[369] William Hagman, Arvid Erlandsson, Stephan Dickert, Gustav Tinghög, and Daniel Västfjäll. "The effect of paternalistic alternatives on attitudes toward default nudges". In: *Behavioural Public Policy* 6.1 (2022), pp. 95–118.

[370] Andrea Larney, Amanda Rotella, and Pat Barclay. "Stake size effects in ultimatum game and dictator game offers: A meta-analysis". In: *Organizational Behavior and Human Decision Processes* 151 (2019), pp. 61–72.

[371] Than Htut Soe, Oda Elise Nordberg, Frode Guribye, and Marija Slavkovik. "Circumvention by design-dark patterns in cookie consent for online news outlets". In: *Proceedings of the 11th nordic conference on human-computer interaction: Shaping experiences, shaping society*. 2020, pp. 1–12.

[381] International Committee of the Red Cross (ICRC). *Who We Are*. Website, Accessed on: 24/08/2023. Available at: `https://www.icrc.org/en/who-we-are`. 2023.

[382] Wikimedia Foundation. *International Red Cross and Red Crescent Movement*. Website, Accessed on: 24/08/2023. Available at: `https://en.wikipedia.org/wiki/International_Red_Cross_and_Red_Crescent_Movement`. 2023.

[383] William Revelle. *psych: Procedures for Psychological, Psychometric, and Personality Research*. R package version 2.3.9. Northwestern University. Evanston, Illinois, 2023. URL: `https://CRAN.R-project.org/package=psych`.

6. Main Contributions, Future Research and Closing Remarks

[372] Don Norman. *The Design of Everyday Things*. Basic Books, 2013.

[373] Vania Dimitrova and Antonija Mitrovic. "Choice architecture for nudges to support constructive learning in active video watching". In: *International Journal of Artificial Intelligence in Education* 32.4 (2022), pp. 892–930.

[374] Armindokht H Sadeghian and Ali Otarkhani. "Data-driven digital nudging: a systematic literature review and future agenda". In: *Behaviour & Information Technology* (2023), pp. 1–29.

[375] Tobias Mettler and Stefan Stepanovic. "Acceptable nudge strategies to incentivize the use of wearables and physiolytics at work: A Q-methodology examination". In: *Journal of Information Technology* (2023), pp. 1–27.

[376] Christine Murad, Cosmin Munteanu, Benjamin R Cowan, and Leigh Clark. "Revolution or evolution? Speech interaction and HCI design guidelines". In: *IEEE Pervasive Computing* 18.2 (2019), pp. 33–45.

[377] Garfinkle, Alexandra. *Amazon has sold more than 500 million Alexa-enabled devices, drops 4 new Echo products*. Website, Accessed on: 02/12/2023. Available at: `https://finance.yahoo.com/news/amazon-has-sold-more-than-500-million-alexa-enabled-devices-drops-4-new-echo-products-140013808.html`, 2023.

[378] Joseph Kupfer. "Privacy, autonomy, and self-concept". In: *American Philosophical Quarterly* 24.1 (1987), pp. 81–89.

[379] Helen Nissenbaum. "A contextual approach to privacy online". In: *Daedalus* 140.4 (2011), pp. 32–48.

[380] Alessandro Acquisti, Laura Brandimarte, and Jeff Hancock. "How privacy's past may shape its future". In: *Science* 375.6578 (2022), pp. 270–272.

8 Appendices

The following sections contain some extraneous information that has been referenced in the previous chapters of the dissertation.

Contents

8.1 Variables in multiple linear regression models

Table 8.1 describes the variables used in the multiple regression models in Chapter 4.

Table 8.1: Variable descriptions in order of appearance for all three regression models featured in Table 4.3.

Variable	Description
$\log_e(\frac{y_1}{y_0})$:	Natural log difference in time spent on site/app between the treatment and baseline periods
d_1:	Strict diet intervention as a dummy variable
d_2:	Self-regulated diet intervention as a dummy variable
$\log_e(x_1)$:	Natural logarithm of the daily average time spent on site/app during baseline period
x_2:	Standardized compulsive use before the baseline period
d_1x_2:	Interaction between standardized compulsive use before the baseline period and strict diet dummy
d_2x_2:	Interaction between standardized compulsive use before the baseline period and self-regulated diet dummy
e:	Residual

8.2 Introductory information about the charity

Figure 8.1 shows the introductory information [381, 382] users received about the charity organization in Chapter 5.

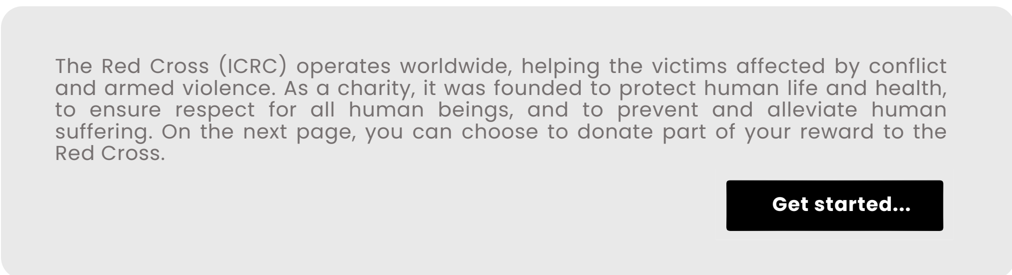

Figure 8.1: The information screen users saw prior to making their donations.

8.3 Distribution and associations between measures

Figure 8.2 shows the associations between the continuous measures of the study in Chapter 5 by combining scatter plots, histograms and Spearman correlations. Figure 8.2 reveals that several measures (like the donation amounts, threats to autonomy, and sense of manipulation) are bimodal as well as having distinct floors. The nature and distribution of these three measures make them ill-suited for parametric statistical test procedures. The latter two measures are also highly correlated. The black dots in the scatter plots highlight the users that changed their donation at the end of the study.

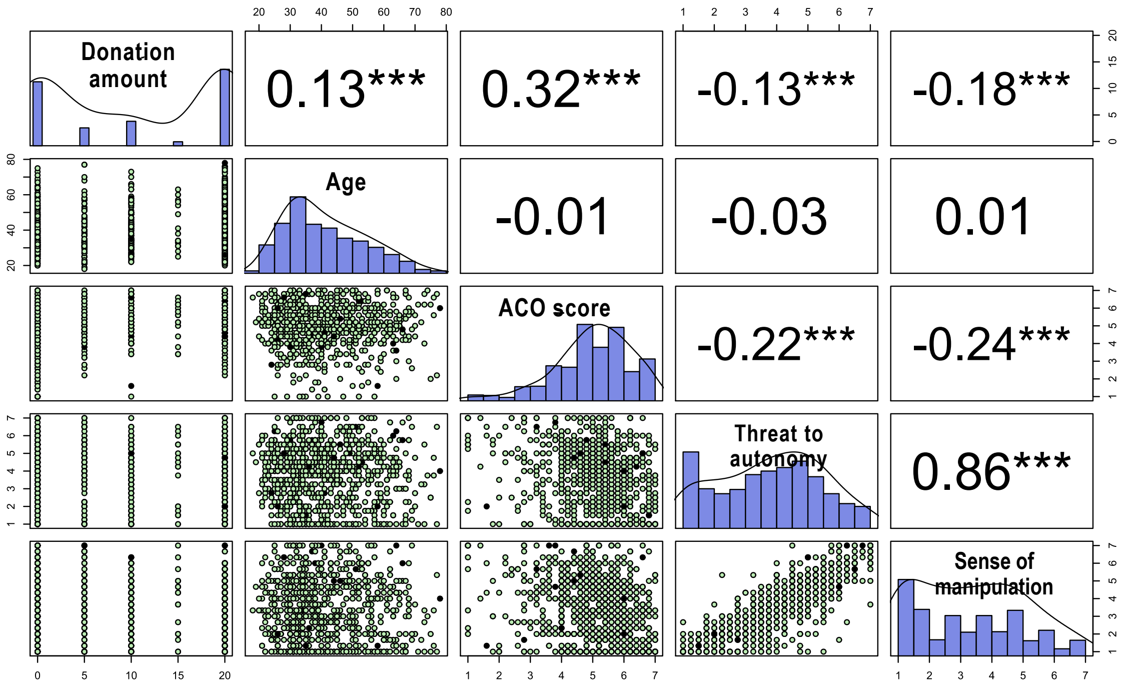

Figure 8.2: Psych plots [383] combining multiple scatter plots, histograms and pairwise Spearman correlations of the variables: donation amount, age, and ACO score, threat to autonomy, and sense of manipulation, ($n = 794$).

8.4 Multiple logistic regression model diagnostics

Table 8.2 shows the model statistics related to the three multiple logistic regression models featured in Table 5.2 in Chapter 5.

Table 8.2: Diagnostics for the three multiple logistic regression models in Table 5.2.

Multiple logistic regression models	-2 Log likelihood	Nagelkerke R^2	†ACC (%)
Model 1: $\chi^2(5, n = 794)$ = 99.766***	974.691	.159	63.7
Model 2a $\chi^2(9, n = 794)$ = 128.183***	946.243	.201	67.4
Model 2b $\chi^2(9, n = 794)$ = 128.183***	946.243	.201	67.4

* $p < 0.05$, ** $p < 0.005$, *** $p < 0.001$
†The classification accuracy (ACC) of naive prediction with zero compliance is 59.4%

8.5 Tests for threats to autonomy, sense of manipulation and compliance

Figure 8.3 further highlights the distribution of the data used in Figure 5.4 together with the results of eight Mann-Whitney U tests. When a Bonferroni correction is applied ($\alpha_E/8 = .00625$) to these Mann-Whitney U tests in Figure 8.3, the same three suggestive differences are detected regarding users' threats to autonomy for the nudge patterns: friction, default+friction, and default+friction+social between those that complied compared to those that did not, see p-values in Figure 8.3.

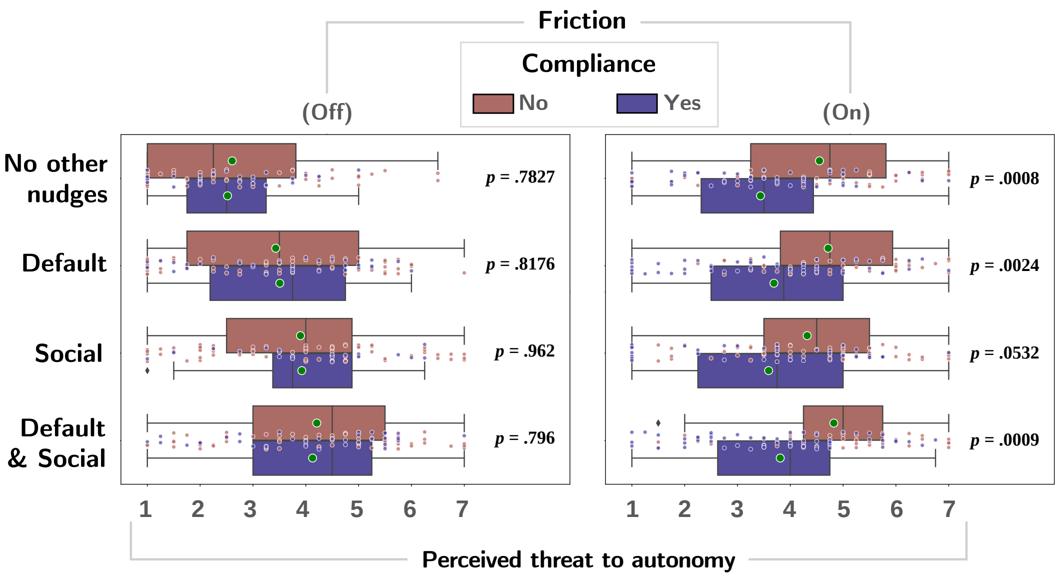

Figure 8.3: Boxplots and scattered rain of the threat to autonomy scale for each of the eight experimental groups, with Mann-Whitney U tests comparing those that did not comply (No) vs. those that did (Yes) within each nudge pattern, ($n = 794$).

Figure 8.4 shows the distribution of the data used in Figure 5.5 together with the results of the eight Mann-Whitney U tests. When a Bonferroni correction is applied to the Mann-Whitney U tests in here, only one suggestive difference remains for users' sense of manipulation between those that complied compared to those that did not. This difference is detected for the friction nudge alone ($.0042 < 0.00625$).

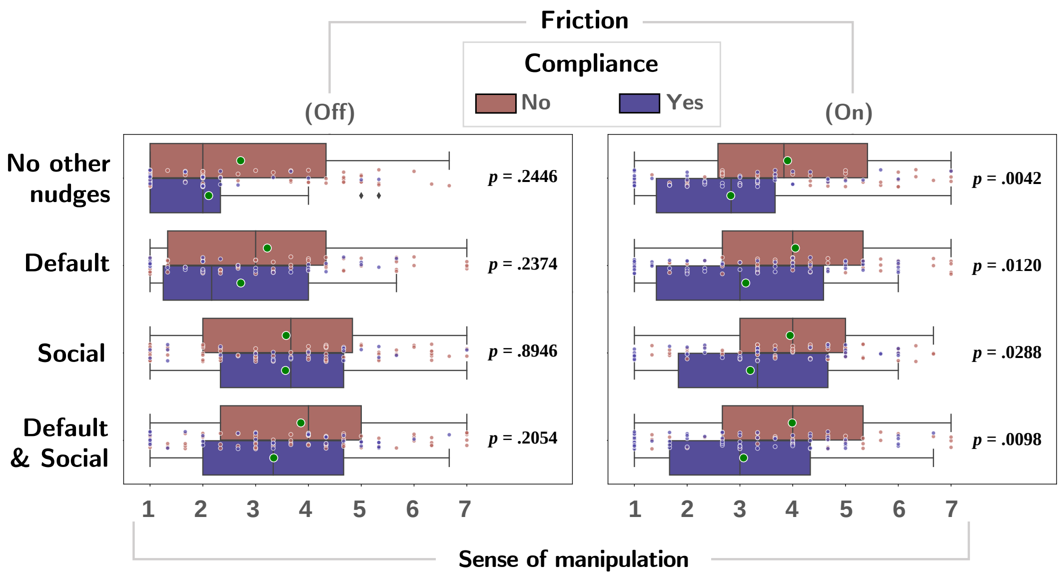

Figure 8.4: Boxplots and scattered rain of sense of manipulation for each of the eight nudge patterns, with Mann-Whitney U tests comparing those that did not comply (No) vs. those that did (Yes) within each nudge pattern, ($n = 794$).