Dark Patterns of digital platforms - when are users fooled by nudging?

Master thesis

Author: Stefan Oberhofer
Matriculation number (Passau): 104172
Matriculation number (Turku): 2104610

University of Passau
School of Business, Economics and Information Systems
Business Information Systems

Turun yliopisto (University of Turku)
Turku School of Economics
Global IT-Management

Supervisor: Sebastian Overhage
Co-Supervisor: Kai Kimppa

September 30, 2022

Declaration of originality

I confirm that the submitted thesis paper is original work and was written by me without further assistance. Appropriate credit has been given where reference has been made to the work of others.
The thesis was not examined before, nor has it been published. The submitted electronic version of the thesis matches the printed version.

Erklärung zur Masterarbeit

Ich versichere hiermit wahrheitsgemäß, die Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung der Universität Passau zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet zu haben.
Diese Arbeit habe ich in gleicher oder ähnlicher Form bisher keiner anderen Prüfungsbehörde vorgelegt.

Passau, den 30 09 2022 .....

Stefan Oberhofer

Contents

List of Abbreviations

2FA Two-factor authentication

AIS Association for Information System

FT Financial Time

OLP Online labor platform

P2P Peer-to-peer

VHB Verband der Hochschullehrer für Betriebswirtschaft

VPN Virtual Private Network

UX User experience

List of Figures

List of Tables

1 Introduction and Motivation

Digital products and services are nowadays a ubiquitous part of our daily lives, act as some guiding tool, provide a connection with others or perform mundane tasks of our highly individualized everyday life. With sophisticated marketing and advertising strategies, digital platforms can offer their customers extraordinarily well-customized and bespoke products (Luguri & Strahilevitz, 2021). These offerings are possible with the help of tracking customer behavior through distributed and collaborative systems (Gray et al., 2021). Furthermore, these companies place the maximization of business and shareholder value in the center by acquiring their customers at the lowest possible cost (Bohnsack & Liesner, 2019). However, from a company perspective, using the customer's data leads to higher productivity levels also exposes customers to manipulative strategies (Gray et al., 2021). In the limelight of digital services and products, the pervasiveness of manipulative practices is on the constant rise, combined with blatantly devious and persuasive design patterns that offer cheap tools for companies to create and sustain their customer base (Gray et al., 2018). Nevertheless, this trend also conducted the emergence of *Dark Patterns*, deceptive design practices exploiting cognitive biases and steering users towards decisions that might affect them negatively. One of the gloomy examples in the industry is Uber, one of the biggest peer-to-peer (P2P) ride-sharing services worldwide, which enables trusted transactions between strangers via a digital platform (Calo & Rosenblat, 2017). The prerequisites for becoming a driver include having a valid driver's license and a good-looking car. After Uber's checking procedure, the new driver works as an independent contractor, not essentially as an employee, apparently giving them all the freedom and flexibility in terms of time, location, and frequency of their work (Lee et al., 2015). Moreover, all the managerial functions familiar in traditional work settings, like work assignments, performance evaluation, and informational support, are carried out only through a smartphone app. The driver has to log into the app to start working, and an algorithm will assign the work tasks (Möhlmann et al., 2021). However, what makes Uber a precedent for Dark Patterns in question?

Imagine the situation of a Uber driver after a long working day, when the driver wants to log off the smartphone app. Usually, the driver expects to get logged off. However, a pop-up message with the headline "Make it to '$40'" and an image with an engine gauge and a needle tantalizingly close to the $40 goal, but was still incomplete, appears (Scheiber, 2017). Below is a message with a text: "You're $6 away from making $40 in net earnings. Are you sure you want to go offline?" (Scheiber, 2017, p. 4). To decide this, the driver has two options: a "Go offline" or a "Keep driving" button, whereas the second button is put in a highlighted contrast (This example can be found in Appendix A). This message pushes the drivers into longer working hours to reach an arbitrary sum of today's earnings.

This example provides insights into the nature of Dark Patterns, a highly prevalent phenomenon in digital platforms (Möhlmann et al., 2021). Through deceptive design techniques, user's choice architecture is altered so that they are being nudged into decisions that benefit the platform owners and not the users (Luguri & Strahilevitz, 2021). But most importantly, data is the fuel in the digital world (Palmer, 2006; Rotella, 2012), getting as much information about the user as possible to offer even more customized products and experiences is crucial for digital platform businesses (Möhlmann et al., 2021). The usage of Dark Patterns also sparks ethical implications in terms of user experience (UX) design and the behavior of platform owners employing them (Gray et al., 2018). In the current literature, the two research streams of digital platforms and behavioral nudging with the help of Dark Patterns are only in a loose connection with each other (Gray et al., 2018; Möhlmann et al., 2021). The first stream offers a profound understanding of the omnipresent construct of digital platforms and its underlying vital concepts, like network effects, information asymmetries, or platform governance (De Reuver et al., 2018; Parker et al., 2016; Tiwana, 2013). Among platform scholars, there has been less previous evidence for the influence of behavioral nudging. The phenomenon of Dark Patterns has been considered more recent (Berger et al., 2022; Özdemir, 2020). Most of the Dark Pattern theories have predominantly been descriptive, objectively categorizing instances of Dark Patterns (Mathur et al., 2021). Due to that reason, a set of related contributions exists but currently lack a particular concern in the perspective of digital platforms and insufficient ethical considerations of Dark Patterns (Gray et al., 2018). Moreover, Maier and Harr (2020) call for a more fine-grained analysis of end-user perceptions concerning Dark Patterns in digital services. Bongard-Blanchy et al. (2021) further argued that future work should also consider less tech-savvy people in end-user accounts related to Dark Patterns.

This paper intends to increase awareness and perception of the Dark Pattern phenomenon by exploring how end-users deal with these patterns and considering the previous implications from other researchers. After defining gaps in the literature, the thesis' objectives are described. The two main theoretical themes of the thesis are only loosely connected. Thus, the first goal of this work is to bridge the two perspectives by conceptualizing the theme of Dark Patterns even further, enhance the theoretical understanding of behavioral nudging on digital platforms and shed more light on the ethical implications of Dark Patterns. This is done by characterizing different aspects of Dark Patterns in the context of digital platforms. Furthermore, the second aim of this work is to investigate the end-user accounts of Dark Patterns by conducting interviews with a broader participant set to address the shortcomings, as mentioned earlier, of technical inexperience among the participants. In general, the following research questions will be addressed by this thesis.

1. What kind of Dark Patterns are used in digital platforms?
2. How do Dark Patterns affect users of digital platforms?

Research question one will be answered by examining previous related Dark Patterns literature and proposing characteristics of Dark Patterns that can be found on digital platforms. The main contribution of the literature is to put the lens on the Dark Patterns onto digital platforms to get insights into what instances of Dark Patterns are used and what characteristics these Dark Patterns exhibit. The present work has significant benefits for further studies to get a more fine-grained understanding of Dark Patterns in digital platforms. Besides, research question two aims to find a solution for end-user accounts of less technically savvy people, widens the understanding of possible user perceptions of Dark Patterns, and indicates practical coping strategies. Research question two will be addressed by using a qualitative approach in the form of semi-structured interviews with individual interviews and focus group discussions. Compared to quantitative techniques, this method might unravel insights into the awareness of the Dark Patterns phenomenon regarding the ethical considerations of end-users.

The remainder of the thesis is organized as follows. It begins by providing theoretical clarity of both perspectives of the thesis context by elucidating platform theory, characteristics of digital platforms, and Dark Patterns. The latter requires an in-depth understanding of the human decision-making process. Therefore, some cognitive science aspects will also be included in this part. A short overview of used ethical theories for the ethical analysis concludes the theoretical background section. Next, a structured literature review helps to identify relevant instances of Dark Patterns in the context of digital platforms. Based on these results, semi-structured interviews with a more diverse participant set, as requested by Bongard-Blanchy et al. (2021) and Maier and Harr (2020) shed light on the awareness and feelings of users exposed to Dark Patterns. In the end, the results will be discussed and interpreted together by pointing out the limitations of this study with possible future research implications.

2 Theoretical Background

This section aims to provide the necessary theoretical framework in the context of the master thesis. Thus, to approach the two perspectives of digital platforms and Dark Patterns, the analysis will be handled in specific subsections. Section 2.1 will conceptualize platforms in general and then emphasize digital platforms with all their related core concepts. Section 2.2 will focus on behavioral science aspects necessary to investigate the phenomenon of Dark Patterns. Lastly, several ethical theories used for the ethical assessment are described in section 2.3.

2.1 Platform Theory

As this paper's primary focus is digital platforms, the following paragraph will outline the most crucial concepts of non-digital platforms and swiftly shift the focus to the scope of this thesis. The objective of this chapter is also to derive a working definition for digital platforms to position the term with its related assignment in the theoretical equilibrium. Over recent years, digital platforms have become a ubiquitous part of our daily life. Today multiple different types of digital platforms exist. For instance, social media platforms like Twitter or Facebook reshaped social interaction between people, and payment platforms like PayPal or Apple Pay transformed the entire financial industry. Disruptive technologies like Data Mining, Artificial Intelligence, or Machine Learning also paved the way for P2P platforms like Uber, which can be summarized under the umbrella term of *sharing economy* or *gig economy* (De Reuver et al., 2018). It is worth noting that platforms are already an existing pattern, even without the need for digital technologies.

Generally speaking, a platform is a particular market and serves as an exchange facilitator between consumers. Without a platform, it would be difficult for them to interact with each other (Gawer, 2014). Furthermore, a platform acts as the coordinator for high-quality match-making of supply and demand (Parker et al., 2016). The supply and demand side of platforms involve different user groups. Therefore, such a multi-sided platform must provide some mediation between these groups (Boudreau & Hagiu, 2009; De Reuver et al., 2018). In this setting, the value of one group often depends on the number of members on the other market side. This phenomenon sparked interest among researchers and is regarded as *network effects* or *network externalities* (De Reuver et al., 2018). Network effects state that the versatility of technology rises when the number of users (so-called installed base) increases (Arthur, 1989). The current literature distinguishes between direct and indirect network externalities. In contrast, the first is about the platform value dependence on the number of users, and the latter refers to the user quantity of a different user group as the main driver for the platform's value (De Reuver et al., 2018). However, indirect network effects can also be of a negative nature - when an increasing amount of users leads to a decrease in the platform's value (Farrell & Klemperer, 2007).

2.1.1 Digital Platforms

The aspects mentioned above about non-digital platforms do not fit perfectly in the pattern of digital platforms, as those are different in terms of the underlying digital technologies (Kallinikos et al., 2013). These technologies entail the homogenization of data, editability, re-programmability, distributedness, and self-referentiality (Yoo et al., 2010). Due to this inherent complexity, there cannot be a single platform owner controlling the platform core and governing design principles in such settings (Hen-

fridsson et al., 2014). Moreover, according to Yoo et al. (2010), combining physical products with the layered architecture of software results in loosely coupled architectures through standardized interfaces. Tiwana and Konsynski (2010, p. 677) describe architecture as the "conceptual blueprint that describes how the ecosystem is partitioned into a relatively stable platform and a complementary set of modules that are encouraged to vary, and the design rules binding both."

Digital platforms can be a challenging field of study due to their distributed nature (Henfridsson et al., 2014). Nevertheless, by conceptualizing digital platforms, there exist three main research streams among platform scholars. The first perspective summarizes digital platforms from an *economic* point of view as markets that have the disruptive potential to change traditional markets and create efficient interactions between consumers and producers with ease (Gawer, 2014; Tiwana, 2013). In other words, a digital platform has a unique value-creation mechanism that enables synergies between external producers and consumers due to digital resources (Constantiou & Kallinikos, 2015). Therefore, the value creation dimension of the underlying business model should be directed towards creating trust and helping the users to carry out efficient and favorable transactions (Bakos, 1998). This trust is formed chiefly by showing the users reviews of previous transactions (Pavlou & Dimoka, 2006). Regarding how value is delivered, the digital platform business model should consider the user's monetary, emotional, and social value (Täuscher & Laudien, 2018). Besides, a digital platform has to capture the value to survive and become profitable. Still, as many *Freemium* services prevail in the digital economy, platforms must consider different means to monetize their services (Täuscher & Laudien, 2018). Abdelkafi and Täuscher (2016) mention revenue streams such as commissions, subscriptions, advertising, or service sales. In this setting, Dark Patterns consider another monetization technique that would willingly influence the user's behavior with the help of human science aspects incorporated into design interfaces (Gray et al., 2018).

Boudreau (2012) and Tiwana and Konsynski (2010) use a more technical lens, as they define digital platforms as purely technical artifacts, where the platform itself consists of an extensible codebase that is complemented by third-party modules integrated into the ecosystem. Third-party developers create these modules (De Reuver et al., 2018). Ghazawneh and Henfridsson (2015, p. 199) extend the definition of Tiwana and Konsynski (2010) and conclude that digital platforms are "software-based external platforms consisting of the extensible codebase of a software-based system that provides core functionality shared by the modules that incorporate with it and the interfaces through which they interoperate." Based on these aspects, a digital platform can also be inferred using a more engineering-centered view. It can be described as technical artifacts with a modular architecture consisting of a stable core and many changing components in its periphery (De Reuver et al., 2018).

However, the last concept of digital platforms - having a solid core that will be extended with various sub-modules created by people from the other side of the platform - entails the necessity of a mechanism to balance the different impulses of the platform users. In the literature, this is referred to as platform governance, which should appropriately balance the distinct interests of all sides in the multi-sided markets of those platform ecosystems (Wareham et al., 2014). *Governance* can be defined as the way how organizations are managed at the highest level and involves, in the context of digital platforms, the division of decision rights between platform owner and third-party developer, the control mechanisms employed by the platform owner, and lastly, incentive structures (Tiwana, 2013). Altogether, these aspects can be grounded in the *organizational* perspective of digital platforms, which bridges the economic and technical views and emphasizes platform development and ecosystem management. According to this perspective, digital platforms are evolving organizations or meta-organizations with three distinct features. First, digital platforms coordinate constitutive and innovative agents, who are in direct competition with each other. Second, digital platforms create value by generating economies of scope in supply and demand. Lastly, they are fostered by a modular and composed technical architecture consisting of a stable core and periphery (Rolland et al., 2018). Generally speaking, a digital platform might not hold the traditional ownership-centric views (Iansiti & Levien, 2004), as they spark dynamics where distributed actors jointly enhance the boundary resources (De Reuver et al., 2018; Eaton et al., 2015).

To some extent, the theoretical platform construct should not just be examined by only considering the technical perspective but also emphasizing social aspects, as a platform often represents some organization with standards and processes. Therefore, this thesis defines a digital platform as an ecosystem built upon a flexible and extendable code-base (through third-party modules), allowing an organizational structure to create and sustain business value.

2.1.2 Gig Economy

After describing digital platforms in detail, this section will present a brief overview of the particular digital platforms used in this thesis. Nowadays, platform-mediated gig work is prevalent in various areas and sectors of our daily and economic life, ranging from taxi drivers to pizza deliverers (Woodcock & Graham, 2019). Gig work can be regarded as a new trend, as it emerged about one decade ago and is now an increasing subject of research for many scholars (Kellogg et al., 2020; Lee et al., 2015; Möhlmann et al., 2021; Vallas & Schor, 2020). Due to this propagation, this paper focuses on online labor platforms in the analysis corpus, especially with more emphasis on the *gig economy*. This section provides a theoretical consensus about these terms by explaining those phrases in more detail.

When encoutering the phrase gig economy, one may initially think about different musicians performing music gigs for their living. This kind of business is merely short-term based, and after a gig, there is no guarantee that the musician can perform on the same spot regularly (Woodcock & Graham, 2019). However, based on the musician's performance, there might be chances to repeat the engagement if the musician did extraordinarily well, is particularly popular, or is just a one-off opportunity. There is also a diverse payment for the musicians, as it can be either a fixed fee, some shares of the ticket entrance fee, or others - in total, their expenses might be covered or not (Woodcock & Graham, 2019). Subject to these aspects of work, several parallels can be drawn to the work discussed in the Uber case introduction. Work tasks in the gig economy are also of short duration and temporary. Therefore, these tasks tend to be precarious and unpredictable (Vallas & Schor, 2020). To receive more work tasks, the worker must be reliable, perform well, and maintain a good reputation (Woodcock & Graham, 2019). The customer should be satisfied with the provided work to rephrase this notion. Furthermore, there are little or hardly any chances of career advancement in many gig work settings (Vallas & Schor, 2020).

Based on the aspects mentioned earlier, the gig economy describes an economic disruption spanning multiple sectors where work is becoming more temporary, unstable, and patch-worked (Woodcock & Graham, 2019). Workers spend less time with just one job, they perform multiple, sometimes also in parallel, and there is a steady risk of time without income even while searching for new employment (Woodcock & Graham, 2019). By way of explanation, the gig economy can be described as labor markets with independent contractors arranged on a digital platform with non-permanent work, variable working hours, little job security, payment on a work basis, and a lack of career development. This work arrangement takes place on online labor platforms online labor platform (OLP).

2.1.3 Online Labor Platforms

How does an OLP fit into the picture of digital platforms? The trend of working on OLP has steadily risen over the years; recent figures showed that 163 million temporarily performing and independent contractors and freelancers in Europe and the U.S. have flexible job arrangements (Manyika et al., 2018). To overcome job insecurity, those people tend to use OLP to secure their work assignments (Möhlmann et al., 2021). Even though the term *platform* is an inherent part of an OLP, current literature focuses more on defining an OLP as a market that enables an efficient matching of supply and demand through machine-learning algorithms (Agrawal et al., 2018; Curchod et al., 2020). Besides the matching perspective, an OLP further provides a model for organizing work, especially by using algorithms to monitor and control the platform workers (Kellogg et al., 2020; Lee et al., 2015; Möhlmann et al., 2021).

These platforms supply the platform worker with a flexible and also autonomous work environment (Deng et al., 2016; Kuhn & Maleki, 2017) enabling surveillance and supervision through the platform owner (Galliers et al., 2017; Newell & Marabelli, 2015). In turn, such OLPs can be considered technology-enabled meta-organizations (being in line with the organizational perspective described in chapter 2.1.1). Their primary purpose is not solely matching supply and demand. Instead, it steps further as it enforces the platform workers towards specific organizational goals (Gawer, 2014). This perspective can be seen in traditional work settings, where control is used to direct and make employees adhere to corporate strategies and rules. A meta-organization "solve[s] the problem of organizing without explicitly relying on formal authority as enshrined in an employment contract" (Gulati et al., 2012, p. 580). Such platform companies combine both the organizational and market mechanism (Constantiou & Kallinikos, 2015), and they have a relatively small platform core but a large periphery (Gulati & Kletter, 2005). For Uber, this notion applies extraordinarily well, as the company has a small core - Uber itself does not own any physical resources like cars, which are only the drivers' property (Möhlmann et al., 2021). Furthermore, Uber does not officially employ human resources like drivers, as the drivers are independent business owners rather than traditional employees (Scheiber, 2017).

However, in this context, a dilemma becomes apparent, as platform companies like Uber can tremendously cut down their labor costs. Still, they cannot compel their workforce to show up at a specific place and time. This lack of control is disproportionate to providing transportation for passengers independent of time and location seamlessly (Scheiber, 2017). To overcome this deficiency, Uber employs decision-making using algorithms to coordinate the drivers to offer highly standardized and valuable services to their customers (Möhlmann et al., 2021). In this setting, technology's importance must also be considered. The aspects of coordination and surveillance mentioned above are only possible with a large-scale data collection used to train algorithms responsible for coordination and control tasks former being carried out by traditional managers (Möhlmann et al., 2021). The surveillance is made possible by analyzing the worker's behavior and performance through algorithms while they carry out their work tasks. In their study, Möhlmann et al. (2021) concluded that one way to enforce algorithmic control is through *Behavioral nudging*, e.g., by altering the behavior of platform workers using algorithms without force in a subtle way and any formal mechanisms like incentive payments. Nudging implies that the users are not forced to undertake any action, e.g., the threat of being banned from the application. However, there are also ways to compel the user into specific actions. Recently, the phenomenon of Dark Patterns - deceptive design interfaces using psychological factors to nudge users into actions not for their benefit (Maier & Harr, 2020) - can also alter the behavior of platform workers. Here it raises the question of how Dark Patterns differ from behavioral nudging. The thesis will continue to

conceptualize the phenomenon of Dark Patterns in the next section to provide an answer to this question.

2.2 Dark Patterns Theory

The services of digital platforms should usually be crafted by following user-centered design principles, meaning that the interface should be designed based on the users' expectations and accommodate their desires (Gray et al., 2018; Moen et al., 2018). The concept of nudging can help designers subtlety nudge users to make decisions by exploiting some cognitive biases the users are mostly not even aware of (Maier & Harr, 2020). Being in line with the practical example listed in the introduction, Uber could show the current earnings in the smartphone app so that the driver can always see, plan their work shifts accordingly and decide when to quit for the present working day. The app can highlight milestone earnings and nudge drivers into driving longer. However, with this construct in place, they are not forced to drive longer - that is the idea of a nudge (Kahneman, 2011). Nonetheless, the psychology behind this example can also be exploited in a more fallacious way, as the driver could be directed into actions that are solely beneficial to Uber. As outlined in the introduction, a pop-up window during the log out of the smartphone app nudges the driver into working longer, which contradicts their initial desire to quit working for the day. It is also ethically debatable, as the user value (Uber driver) is supplanted in favor of the shareholder (Uber) value (Gray et al., 2018). To reach common ground, these deceptive design techniques, implementing functionalities, not in the users interest and deliberately misleading them, are commonly referred to as Dark Patterns in the literature (Gray et al., 2018; Maier & Harr, 2020). Using this derivation, Dark Patterns exploit the concept of nudging in a more deceptive and elusive way. Nevertheless, careful consideration of the human decision-making process is inevitable to deploy Dark Patterns. These design interfaces can be crafted after the predictable human behavior to either adhere to or change the behavior (Maier & Harr, 2020). The underlying mechanism of Dark Patterns and nudging will be explained further to put these concepts in the appropriate theoretical context. The following section briefly introduces the human decision-making process and specific cognitive functions, followed by the concept of nudging and digital nudging. Lastly, the subject of Dark Patterns and a conceptualization of its nature will be presented.

2.2.1 Human Decision-making Process

Dark Patterns can influence a person's decisions and alter their behavior. Nevertheless, this contradicts the classical management theory that describes humans as rational decision-makers or *Homo Economicus* (Koontz, 1961). However, people tend to be prone to errors and systematically deviate from rational decision-making,

putting the classical claim at risk of withstanding further on in the context of Dark Patterns (Özdemir, 2020). The way Dark Patterns work requires an in-depth understanding of decision-making, as human behavior can mainly be predicted. Based on this knowledge, user experiences are crafted to either support or alter humans' behavior (Maier & Harr, 2020). The central premise of cognitive psychology among scholars states that two different systems are actively involved in human decision-making (Weinmann et al., 2016). Kahneman (2011) built upon the *Dual process theory* of Wason and Evans (1974) and labeled these two systems as *System 1* and *System 2*. *System 1* works in an unconscious, fast and effortless manner and uses emotions and intelligible heuristics to make decisions (Maier & Harr, 2020). As a result, the decision outcomes create predictable biases that could harm human beings (Tversky & Kahneman, 1974). *System 2* performs decision-making much slower and more conscious compared to *System 1*, as it "construct[s] thoughts in an orderly series of steps" (Kahneman, 2011, p. 21). Moreover, it uses the experience of concentration and choice. These two systems generally serve different purposes and have distinct features and limitations. As the human brain capacity is limited, no human can observe and process everything that is around somebodies instantaneously changing environment. Therefore, about 95 % of all cognitive actions happen in a non-conscious manner, as the human brain could not handle and process this constant and enormous amount of data inflow (Raichle, 2010). Still, to provide feedback, Kahneman (2011) suggests that *System 1* responds intuitively, whereas *System 2* oversees the quality of those made responses - some of them might be endorsed, corrected, or overwritten.

As indicated previously, *System 1* heavily relies on heuristics and biases in making decisions. A heuristic is regarded as a mental shortcut used to judge or solve problems, like gut feelings or listening to social conformity (Kahneman, 2011; Maier & Harr, 2020). Yet, relying solely on heuristics does not guarantee to be the silver bullet for making decisions. Having extensive confidence in heuristics when making decisions make humans vulnerable to a logical and cognitive fallacy called *cognitive bias* (Maier & Harr, 2020). At this point, there are two possible directions in how these biases can be used. On the one hand, design interfaces can use the biases to help people improve their decisions. On the other hand, biases can also be adapted to steer users into undesired or malicious intents (Fischhoff, 2002), which is the core nature of Dark Patterns. Seminal contributions have provided an overview of certain cognitive biases exploited through Dark Patterns. The following list collects the most important biases and briefly explains their functionality.

- *Anchoring Effect*: Users tend to rely on the first piece of the received information in subsequent decisions (Kim et al., 2021; Tversky & Kahneman, 1974)
- *Bandwagon Effect*: Individuals value something more due to peers valuing the same (Sherif, 1936)

- *Default Effect*: Situations where preferred choices are preselected (Thaler & Sunstein, 2009) and users stick with default provided options due to inertia (Johnson et al., 2002)
- *Framing Effect*: Information is presented differently to individuals, who then make different decisions (Berger et al., 2022; Tversky & Kahneman, 1985)
- *Scarcity Bias*: Scarce things are being valued more than non-scarce items (Kim et al., 2021)
- *Sunk Cost Fallacy*: Individuals tend to stick to an even worse option if they already invested resources (e.g., time, money, or emotions) into it (Arkes & Ayton, 1999)
- *Hyperbolic Discounting*: Tendency of individuals to select smaller short-term rewards instead of long-term gains (Waldman, 2020)

In addition to heuristics, the mood is also a factor of influence in the decision-making process of humans. However, it is a variable that is highly amenable to being influenced by various determinants (Tversky & Kahneman, 1974). In this way, it is possible to assess whether a decision is being made due to logical reasoning (Shafir et al., 1993) or because of emotions (Bechara, 2004). In this sense, Västfjäll et al. (2016) differentiate between *integral* and *incidental* emotions. *Integral* emotions refer to the internal representation of a decision and its direct influence. On the other hand, incidental emotions can be considered heuristics and a vehicle for making evaluative choices. Importantly, *integral* emotions are controllable; incidental emotions are not and can be specially addressed with the help of nudging.

How do cognitive biases relate to human decision-making and Dark Patterns? Individuals use two modes of thinking, namely *System 1* and *System 2*. *System 1* is fast and works unconsciously, but all processed information lacks specific details and context. It highly depends on the usage of heuristics due to bounded rationality, which relies on cognitive biases. *System 2* is slower, conscious, more rational, and uses reasoning and logical evaluation of evidence (Kim et al., 2021). As *System 1* acts as a pre-processor for *System 2*, impressions and intuitions from *System 2* will be transformed into beliefs or norms. In other words, *System 2* may consider suggestions from its counterpart without any modifications, leading to conclusions and irrational decisions based on cognitive biases (Kahneman, 2011). Thus, these cognitive biases inhibit rational decision-making and can be exploited through nudging or, even worse, Dark Patterns.

2.2.2 Nudging and Digital Nudging

The terms nudging and digital nudging seem to be used interchangeably, and the transitions are fluent in current literature. In general, designers use the aforementioned cognitive biases and emotions to effectively nudge users into particular choices without restricting their freedom of choice (Moen et al., 2018). However, the question of what exactly nudging and digital nudging refer to remains unanswered. A nudge can be described as a subtle change in the choice architecture of individuals that predictably alters their behavior without forbidding any options or changing their economic incentives to a significant extent. In addition, they point out that the central idea of a nudge is to know how people think, and it should be easy and cheap to implement (Thaler & Sunstein, 2009). As indicated earlier, *System 1* utilized heuristics for its quick decision-making process, but it is vulnerable and prone to errors. This finding leads to a decision to the detriment of the human being (Berger et al., 2022). Nudging counteracts this problem and influences people, making the decisions more predictable (Thaler & Sunstein, 2009). The concept of nudging has been considered to be effective across multiple domains. Examples of nudging can be found, e.g., in organ donation (Johnson et al., 2002) or traffic light labels to encourage healthier food (van Epps et al., 2016).

Due to the progressing digitization, making decisions increasingly takes place in online environments (e.g., apps or browsers). The term nudging can be extended with this notion to *digital nudging*. Weinmann et al. (2016, p. 433) describe *digital nudging* as using "user-interface design elements to guide people's choices or influences users' input in online decision environments." Meske and Potthoff (2017) add that the individual's freedom of choice is not restricted in this sense by digital nudges. Weinmann et al. (2016) further concluded that one of the significant strengths of digital nudging is prompt implementation, evaluation, and personalization at low costs. Moreover, it is relatively easy to test multiple designs instead of waiting years to retrieve the effectiveness, and they offer unprecedented scaling potential (Özdemir, 2020). Generally speaking, (digital) nudges alter the choice environment of people to steer them towards certain decisions, but without restricting their choices. On the contrary, the choice architecture can also be changed through nudges in a more fallacious and delusive way so that the benefits are against the user's interests. Due to this dark and evil nature, the distinction between nudges and their darker counterpart, namely Dark Patterns, can be drawn.

2.2.3 Dark Patterns

In the digital platform world, people are typically unaware of deceptive design patterns crafted to hide particular things and to misdirect the users' attention to what a company wants the users to see (Maier & Harr, 2020). Despite the existing

unawareness, the users have most likely faced such instances in their everyday life, e.g., in social media (Moen et al., 2018), video games (Zagal et al., 2013), or mobile applications (Bösch et al., 2016). Users of digital platforms can be directed into actions through dubious design techniques that exploit cognitive biases. The outcome benefits the company to the user's disadvantage (Moen et al., 2018). As this approach also raises ethical concerns, this concept has been named *"Dark Patterns"* by UX practitioners (Gray et al., 2018). One of those, Harry Brignull, defined a Dark Pattern as "a user interface that has been carefully crafted to trick users into doing things...they are not mistakes, they are carefully crafted with a solid understanding of human psychology, and they do not have the user's interests in mind" (Brignull et al., 2010). The term "dark" carries two distinct meanings. Firstly, it refers to the notion that designers are consciously performing something unethical. Secondly, the users are unaware of it (Gray et al., 2018).

However, this definition leaves plenty of means for interpretation. First, it raises the question of whether it is justifiable to blame certain businesses with the imprint of such a negatively connoted phrase when a design element of a digital platform can be poorly designed. This phenomenon is regarded as *Anti-Patterns* but can be considered the minority among deceptive design elements prevailing on digital platforms. Most designs result from clear and purposeful design ideas and intentions (Gray et al., 2018). Di Geronimo et al. (2020) stated that 95 % of all applications they examined contained at least one form of Dark Patterns. Furthermore, the question about what the user is being tricked into and motivational aspects remain unclear considering the definition above. Moreover, the description above lacks ethical implications about the end-user consequences (Mathur et al., 2019).

The extant literature on Dark Patterns generally provides four facets that are only covered to some minor extent by the definition above. The first facet is about Dark Pattern characteristics having the potential to affect users, as they can be design tricks (Waldman, 2020), misleading interfaces (Bösch et al., 2016) or have "obnoxious, coercive or deceitful" properties (Gray et al., 2018). The second perspective deals with the effect of influencing users, as Dark Patterns are used to trick users (Bösch et al., 2016), subvert user preferences (Luguri & Strahilevitz, 2021) or undermine user autonomy (Maier & Harr, 2020). Third, most definitions provided by scholars neglect the role of the UX designer, as they use and abuse human science aspects to achieve a goal (Conti & Sobiesk, 2010; Gray et al., 2018). Fourth, the question about potential benefits and harms stemming from those user interfaces is only touched to a minor extent by current literature (Gray et al., 2018; Mathur et al., 2019; Waldman, 2020).

The thesis provides a conclusive definition of Dark Patterns to connect the different perspectives. Dark Patterns can be described as misleading user interfaces whose designers use the knowledge of human behavior to trick users into taking specific actions that supplant the user desires in favor of platform provider objectives. This

definition further highlights the ethical consequences for users in a more abstract fashion. This description also conforms with the theoretical developments in the domain of Dark Patterns, which can be divided into ethical UX design and honest end-user perceptions (Gray et al., 2018; Maier & Harr, 2020).

Design is a persuasive act, where the designer can intentionally change the behavior or social norms (Gray et al., 2018; Redström, 2006). However, few studies have conceptualized persuasive technologies and their ethical impacts on individuals over the past two decades. According to Fogg (2009), a persuasive technology shapes the desired behavior through strategies like reduction, tunneling, tailoring, suggestion, self-monitoring, surveillance, and conditioning. Based on that, persuasive design is initially claimed to be for good. However, suppose designers abuse the power of design, e.g., by using countdown timers to create some sense of urgency and pressure among the users. These resulting instances can be classified as Dark Patterns, as they are explicitly designed to persuade or deceive people and pose fierce ethical issues (Gray et al., 2018; Mathur et al., 2019). In other words, Dark Patterns pose the darker version of design interfaces, with explicit persuasive intentions underlying all the design activity (Gray et al., 2018).

By looking at the ethical implications for users exposed to Dark Patterns, several consequences emerge, as outlined in the following course of this thesis. First, Dark Patterns might steer users into actions they would not carry out if they were fully aware and informed about them (Mathur et al., 2019). Consequently, users make choices that are not in their best interest and may deprive their agency. Considering prevailing information asymmetries on digital platforms is ethically even more problematic, as by placing confident trust in the service provider, the user assumes that the service provider knows what is best for them (Moen et al., 2018). The information asymmetry becomes apparent as users in many digital platforms cannot ascertain privacy exposing risks (e.g., digital platforms trade the user data) in combination with a short-term monetary advantage in the form of a one-time discount. Dark Patterns may further annoy or frustrate users, circumvent them (e.g. by financial losses), trick them into sharing more personal data than initially desired, or pushing them into compulsive or addictive behavior (Di Geronimo et al., 2020; Mathur et al., 2019). The digital platform universe consists of different types of Dark Patterns. Chapter 3.2.2 will shed light on the types of these dark instances on digital platforms.

2.3 Ethical Theories

This section provides a brief overview of the theories used for the ethical assessment of Dark Patterns and related end-user accounts in this thesis. The ethical lens will examine raised issues using *non-consequential* (*Kantian Deontology*), *consequential* and *utilitarian* theories. Utilitarianism is a form of Consequentialism and

describes the most ethical choice as the one that produces the greatest good for the most significant number of people (Sinnott-Armstrong, 2003). In other words, for each action, a cost/benefit analysis is carried out to measure the greatest number of good. It is important to note that the resulting good does not necessarily imply happiness (Stahl et al., 2014). However, the utilitarian approach has some pitfalls. It requires predicting the future outcomes of actions to carry out the cost/benefit analysis to determine ethically right or wrong actions. As the utility is challenging to measure ahead of time with certainty, the Consequentialism perspective, in contrast, assesses ethical questions solely based on an action's consequences but avoids direct measurement. Nonetheless, Consequentialism does also not solve the question of predicting the future with certainty. Therefore, it is also difficult to forecast the result of actions ahead of time. Furthermore, ethically sound decisions with a consequential lens can be objectionable, even though the consequences are arguably good. In contrast to Consequentialism, Kant's morality construct deals with acting morally soundly following a universal law and underlying categorical imperatives (Kant, 1870). Duty is a central premise in this setting, to intend something good for another person to be good. According to Kant (1870), the imperatives involve always treating a person as a valuable human being and not merely as a means to an end (*imperative one*), and when it would be a universal (moral) law that all could expect to follow (*imperative two*). Imperative three states that an action can only be good when being voluntary. In other words, actions can be considered ethical if they do not violate the universal law (e.g., no cheating, stealing, or lying). Following the rules makes Deontology easy to apply and does not require weighing the costs and benefits of a situation. Therefore, subjectivity and uncertainty coming from consequences in the future can be avoided by using the Kantian rule-based approach. Despite its strengths, rigidly using the deontological approach can produce unacceptable results by disregarding the possible consequences of actions during determining right and wrong actions. Lastly, Deontology and Consequentialism can be used as a combined approach by looking at the intent and consequences of actions (Kimppa, 2020). According to this notion, an action is morally sound when the intent and consequences are good. Contrarily, when the intent is wrong or even evil, the outcome is ethically not acceptable, notwithstanding any accidentally good consequences, as those do not justify the overall goodness. However, intentions are often hard to know beforehand, making ethical assessments with the combined approach complex (Kimppa, 2020).

3 Structured Literature Review

As of now, the question of what types of Dark Patterns exist in the world of digital platforms remains still unanswered. Therefore, research question one of this thesis will be addressed in this section by conducting a structured literature review.

The remainder of this part is organized as follows. First, the underlying methodology is described thoroughly and results in a concept matrix of the structured literature review. Subsequently, the different concepts of the matrix are discussed in detail.

3.1 Methodology

One of the biggest problems among IS researchers is the lack of explicit methods or established guidelines for documenting the literature search process. In particular, the exclusion and inclusion of specific sources must be completely transparent to prove a review's credibility and exhaustiveness (Brocke et al., 2009). Therefore, the conducted literature review was structured following the recommendations of Brocke et al. (2009) and Webster and Watson (2002). Generally speaking, this procedure involved the scope definition of the structured literature review and elaborating a search string with specific keywords and synonyms to query online databases. Furthermore, it also included the sources (databases and journals) where the review was conducted, and criteria about how the result set was filtered and stripped down (e.g., by title/abstract or content). Finally, from the resulting small final paper set, several concepts that strongly cohere with the review objective are retrieved and collected in a concept matrix.

The structured literature review intended to "Find different types of Dark Patterns." As indicated earlier, the research streams of digital platforms and Dark Patterns are only loosely connected. Therefore, the scope definition of the review did not consider *digital platforms* explicitly. This attempt allowed the collection of a vast and exclusive amount of Dark Pattern types, not restricted to any particular digital platform. Instead, it is one of the thesis' contributions to the literature to connect the dots of the two so far hardly related perspectives.

The search string was created by following the concept mapping idea, which also derived possible relevant search terms like synonyms or other related concepts (Brocke et al., 2009). However, literature can hardly be exhaustive, as likely relevant articles that do not include the selected search terms might not be considered at first glance of the literature search. Furthermore, as both literature streams of digital platforms and Dark Patterns are merely connected, there is no restricting keyword phrase such as *digital platform* included in the search string. This approach allows for a comprehensive result set, and the thesis will provide a narrowed-down perspective of Dark Patterns in digital platforms as its research contribution. The following search terms are combined into the search string and are retrieved per the objective of the literature review: "*dark pattern** **OR** *anti pattern** **OR** *deceptive design* **OR** *evil design* **OR** *dark design* **OR** *ethical design* **OR** *dark ux* **OR** *nudging* **OR** *cognitive bias**." The selection and combination of this keyword string is crucial for the ongoing review process, as it already sets "the parameters of the research itself" (Baker, 2000, p. 222).

After conceptualizing the search term, *EBSCOhost* was queried using the search term and the *Business Source Premier* database, one of the most utilized business research databases worldwide. Using the search string allowed filtering out several articles that were not relevant to the given research question of the thesis. It is worth noting that peer-reviewed articles were put in focus, followed by proceedings of renowned conferences. However, those quality contributions are usually less mature and inferior than journal articles (Brocke et al., 2009). Conference proceedings mainly cover the emerging literature on Dark Patterns. Those articles with higher quality (e.g., citation count on *Google Scholar* or *EBSCOhost*) were used. As the literature review only considers high-quality journals, therefore, the result set is filtered accordingly by applying several rankings (Association for Information System (AIS) Senior Scholar's Basket of Journals, Financial Time (FT), Verband der Hochschullehrer für Betriebswirtschaft (VHB) with ranking of A+ to B). After stripping down the paper set, *forward* and *backward* searches extended the result set. The first technique refers to searching additional literature sources that have cited the particular article retrieved from the keyword search (Brocke et al., 2009). The latter focuses on reviewing older literature cited in the found articles (Webster & Watson, 2002). A three-folded title/abstract/content scanning approach was then used to analyze the paper, always with the literature review's objective to include or exclude articles from the result set. For instance, the article of Kim et al. (2021) has the title: "Dark Patterns used by online travel agencies." Due to the literature review's intention to find types of Dark Patterns in a broader context, the article was included in the result set. This procedure was applied to the entire result set of approximately 750 papers when theoretical saturation occurred so that the literature did not provide new concepts. According to Webster and Watson (2002), a literature review can be considered nearly completed when no new concepts in the literature can be identified. In total, 17 articles were considered to be relevant for answering research question one of the master thesis. Appendix B provides a detailed flow chart of the structured literature review.

The remainder of this thesis requires a concept matrix for the literature review. Therefore, this organizing framework was structured by following a concept-centric approach where key concepts can differentiate between the identified articles. Table 1 briefly illustrates the basic concepts of the articles, a more precise and fine-grained version can be found in Appendix C.

Articles	**Concepts**					
	Dark Pattern Strategies	**Dark Pattern Types**	**Dark Pattern Characteristics**	**Cognitive Biases**	**Effectiveness**	**Usage Motivation**
Baroni et al. (2021)		✗	✗			
Berger et al. (2022)						✗
Bösch et al. (2016)	✗	✗				
Chromik et al. (2019)	✗			✗	✗	
Conti and Sobiesk (2010)	✗	✗		✗		
Di Geronimo et al. (2020)	✗					
Gunawan et al. (2021)		✗	✗			
Gray et al. (2018)	✗	✗				
Kim et al. (2021)		✗				✗
Luguri and Strahilevitz (2021)	✗	✗		✗	✗	✗
Mathur et al. (2019)	✗	✗	✗	✗		✗
Mathur et al. (2021)			✗			
Mildner and Savino (2021)	✗	✗				
Moen et al. (2018)		✗		✗		✗
Özdemir (2020)		✗				✗
Waldman (2020)						✗
Zagal et al. (2013)	✗	✗				

Table 1: Aggregated concept matrix

3.2 Results

3.2.1 Dark Pattern Strategies

Scholars have made seminal contributions to structuring different types of Dark Patterns, which are primarily based on a taxonomy developed by Harry Brignull (Brignull et al., 2010) during the emergence of Dark Patterns around one decade ago (Gray et al., 2018; Luguri & Strahilevitz, 2021; Mathur et al., 2019). Interestingly, even before this period, Conti and Sobiesk (2010) attempted to classify malicious design interfaces but did not label them as Dark Patterns, as this phrased developed some years later. The authors came up with 11 categories, which can further be differentiated into subcategories. They found a category called **Coercion** that menaces the users' compliance, e.g., by requiring them to enter mandatory contact details before accomplishing specific actions. Another instance of malicious design applies to **Confusion**, where users are tangled by asking incomprehensible questions with double or even triple negatives. Furthermore, **Distraction** means detracting the users' attention easily by exploiting their perception, e.g., through advertisements or eye-catching designs. Moreover, users are prone to errors in the digital world, which can be exploited to reach specific goals, e.g., by displaying ads instead of help instructions when the user enters an incorrect URL (**Exploiting Errors**). Putting an extra burden on the cognitive workload of users' actions is considered to be **Forced Work**. It can be seen, for instance, in the uninstalling process of the default Edge browser on Windows devices, which requires doing sophisticated registry edits. The task flow of a user can further be interrupted, e.g., by forcing them to view certain ads before viewing a newspaper article (**Interruption**). In addition, navigation can be manipulated to steer the user towards the desired designer goal, e.g., by asking a mere infinite number of questions to obtain a free electric device (**Manipulating Navigation**). Especially in the context of cookie consent banners, hiding desired information or putting the reject button in reduced contrast may lead to accepting tracking instead of only the technical necessary cookie options (**Obfuscation**). The taxonomy further includes the category of **Restricting Functionality**, where necessary user controls are limited or omitted to reach a specific objective (e.g., no "unselect all" option in a pre-specified email list selection available). **Shocking** a user by displaying controversial or undesired content to the user is further listed among the other types. Lastly, misleading the users by deceiving them through downloading software silently in the background or spoofing content (**Trick**) concludes the list of malicious design interfaces. This exhaustive list provides a good entry point to draw parallels and connect different perspectives with the Dark Pattern strategies of other researchers.

Gray et al. (2018) used the work of Brignull et al. (2010) and defined five primary Dark Pattern strategies serve as a strategic motivator for designers. **Nagging** delin-

eates as a frequent minor redirection of a user from its current task (Di Geronimo et al., 2020; Gray et al., 2018). This Dark Pattern strategy fuses explainability and user control with hidden functionality and steers the user towards unintended actions or interruptions of their tasks (Chromik et al., 2019). This strategy seems to align with the **Obfuscation** and **Interruption** strategies of the taxonomy of Conti and Sobiesk (2010).

The following strategy is called **Obstruction** and refers to any impedance of the users' task flow by exacerbating an interaction more than it inherently needs to be in the form of placing a significant barrier in specific tasks (Di Geronimo et al., 2020; Gray et al., 2018). Users should, in turn, dread the effort to find plausible explanations about the barrier and instead follow the provided solution by the platform owner (Chromik et al., 2019). This category aligns best with the **Forced Work** class of Conti and Sobiesk (2010).

All related hiding, disguising or delaying the divulging of user-relevant information is concerned with the category of **Sneaking** (Di Geronimo et al., 2020; Gray et al., 2018). In this setting, a certain amount of the user's interaction is consumed through hidden functionalities (Chromik et al., 2019). Conti and Sobiesk (2010) describe this category in their corpus as **Coercion**.

Luguri and Strahilevitz (2021) add the strategy of **Interface Interference**, any manipulation of design interfaces to favor specific actions over others to confuse the users. Chromik et al. (2019) amend that this strategy emphasizes certain information or controls that the platform provider prefers. This category coincides best with the **Obfuscation** strategy retrieved from previous work of Conti and Sobiesk (2010).

The last primary class of Dark Pattern strategy refers to **Forced Action** and describes situations where the user has to perform particular (undesired) actions to continue or access any functionality of the digital service (Di Geronimo et al., 2020; Gray et al., 2018). In this context, users are provided with different explanations and action controls so that they are steered towards any desired outcomes by the platform provider (Chromik et al., 2019). The category of **Forced Work** in the corpus of Conti and Sobiesk (2010) provides a similar meaning to this Dark Pattern strategy. So, generally speaking, the taxonomy of Conti and Sobiesk (2010) is in line with the recent work of Gray et al. (2018).

The previously listed Dark Pattern strategies serve as an essential basis among scholars (Chromik et al., 2019; Di Geronimo et al., 2020; Gray et al., 2018; Luguri & Strahilevitz, 2021; Mildner & Savino, 2021). However, this corpus can be extended by considering the work of Bösch et al. (2016), Mathur et al. (2019), and Zagal et al. (2013). Zagal et al. (2013) focused in their article on the identification of Dark Patterns in the gaming context and concluded their analysis with three unique strategies. The first strategy is **temporal**, as the game's players are tricked out of their time, e.g., by performing repetitive and tedious tasks to take more or less time than players

expected. Besides time as the basis, money can also be employed by **monetary** Dark Patterns. Through this strategy, players are deceived into spending more money than they initially anticipated or even paying their money on unexpected occasions. Lastly, **socio-capital-based** Dark Pattern strategies are also quite common among games, which are undoubtedly considered a social activity and take advantage of the players' value of their social standing and relations (social capital). For instance, players feel socially obliged to play and must invite others to make progress in the game.

Besides the gaming context, Bösch et al. (2016) provide a systematic and theoretical analysis of privacy-related Dark Pattern strategies. They mention **maximize** (collecting a disproportionate amount of user data), **publish** (from the public expected to be hidden personal data is disclosed), **centralize** (data processing is carried in a centralized way), and **preserve** (interrelationships between data items are affected by processing, e.g., through aggregation). In addition, **obscure** (users are not capable of informing themselves what happens with their collected data), **deny** (denial of control over the users' data), **violate** (the privacy policy shown to the user is broken by intention), and **fake** (pretense of responsible entity to adhere to strict privacy protection) complete the list of privacy-related strategies. This new perspective allows a more unified view of Dark Patterns, providing some central premises on how user privacy can be exploited.

As indicated earlier, most studies (e.g. (Di Geronimo et al., 2020; Luguri & Strahilevitz, 2021; Maier & Harr, 2020; Mathur et al., 2019)) trying to conceptualize Dark Patterns relied on a list provided by Brignull et al. (2010) that already dates one decade back. Due to recent technological advancements, especially in UX design, Mathur et al. (2019) add some new Dark Pattern strategies to the corpus to sufficiently cover these novel techniques. Firstly, they define an **Urgency** strategy, which refers to evil design interfaces imposing a deadline in the form of countdown timers or similar to nudge users into making quicker (and maybe less thought through) purchases. Secondly, humans tend to rely on the actions and behavior of others to the determination of their behaviors' correctness. This principle can be exploited through the **Social Proof** strategy, leading to users' faster purchases. Lastly, yielding the signal of limited availability or high demand for certain products can increase users' overall perceived value and desirability. The authors denote this strategy **Scarcity**, concluding the list of Dark Pattern strategies.

3.2.2 Dark Pattern Types

After examining the existing Dark Pattern strategies, the underlying instances of Dark Patterns are studied in the remainder of this section. The analysis refers to a created taxonomy based on the identified concepts in the literature review. It can be found in Appendix D. The following instances of Dark Patterns will be aligned

with the specific Dark Pattern strategy discussed in the previous section. *Nagging* was defined earlier as a frequent minor redirection of expected service/product functionality. In the extant literature, no specific instances are described as there exists a varying level of nagging behavior and potential malice that occurs not only randomly (Di Geronimo et al., 2020; Gray et al., 2018). Mathur et al. (2019) refer to *Nagging* as a broader Dark Pattern class that does not cover any subcategories. One example can be seen in the introductory case of Uber, where a pop-up window nudges the driver into driving longer by showing the remaining profit to an arbitrary income sum.

The Dark Pattern strategy of *Obstruction* includes four instances in total and is all about the dissuasion of actions by making it harder for the user to choose them. The **Roach Motel** pattern applies to situations that a user gets into with ease. However, dropping out is almost impossible (Gray et al., 2018; Özdemir, 2020). For instance, the user must call a phone number to cancel their account at a specific service. Furthermore, information about account cancellation is primarily not included in frequently asked question sections on the websites/apps of the services (Di Geronimo et al., 2020). Mathur et al. (2019) also included this type of Dark Pattern in their result set but labeled it as **Hard to Cancel**. **Price Comparison Prevention** exacerbates comparing of prices on digital services by making certain product information (e.g., cost or product-ID) not copyable to hinder users from searching the product on other websites or search engines (Baroni et al., 2021; Özdemir, 2020). Whenever there is some virtual currency in place, the **Intermediate Currency** pattern may occur, as the users are disconnected from the actual value of their currency and thus, have a different expenditure pattern compared to real money (Di Geronimo et al., 2020; Gray et al., 2018; Gunawan et al., 2021). The last pattern in this strategy makes references to **Immortal accounts**, which states that users on digital platforms cannot delete their accounts and personal information (Luguri & Strahilevitz, 2021). This pattern can be expressed by hiding any account deletion options or the pretense of profile erase while still using some data (Bösch et al., 2016). In this context, the **Roach Motel** and **Immortal Account** patterns look similar. However, the first one refers more to performing actions like calling a phone number to cancel the account. In contrast, the latter emphasizes completely missing erase options or the continued use of data despite profile deletion.

Sneaking concerns all user-relevant information disguising, divulging, or hiding actions and can be divided into four different variants. The **Sneak into Basket** pattern can be seen in digital marketplaces when items are added to the shopping cart of the users without their consent, usually being indicated as applicable supplements based on the existing cart items (Di Geronimo et al., 2020; Gray et al., 2018; Özdemir, 2020). If users did not check their cart before checkout, this might trigger the purchase of things that were not the user's initial purchase intention. Also, before checkout, and mainly after tediously adding shipping/billing information, there can be the situation

that additional and often unusually high charges (**Hidden Costs**) are added on top of the checkout price (Kim et al., 2021; Mathur et al., 2019; Özdemir, 2020). With this technique in place, the users mainly justify the additional charges by accepting the purchase and do not consider it a bootless errand (Gray et al., 2018). These additional charges could also be recurring when a user agrees to a subscription that pretends to be a free trial. Usually, the users become aware of the regular transfer at some later point in time, and this **Hidden Subscription** or **Forced Continuity** pattern (Gray et al., 2018; Luguri & Strahilevitz, 2021) comes together with the aforementioned **Roach Motel** pattern (Mathur et al., 2019). The last variant of this strategy is called **Bait and Switch**. In general, users in the digital world have at least a basic understanding of the expected functionality, e.g., when they press a button (Özdemir, 2020). For instance, the current action will be aborted after clicking on a red "X" button or a pop-up window is closed. This pattern comes to light when another activity is performed instead of the desired one. The example can be thus extended by downloading specific software after pressing the red "X" button instead of closing the pop-up window (Gray et al., 2018).

The subsequent strategy, *Interface Interference*, was previously defined as any disturbance of the interface design to favor particular action over others and can be divided into eight instances. First, the **Aesthetic Manipulation** pattern can be found whenever relevant information is not immediately accessible to the user (Gray et al., 2018). In other words, important information is obscured by visuals. It can be seen while registering for a digital service when the accept the terms and conditions button is followed by a small-font-sized text about the data selling practices of the service (Luguri & Strahilevitz, 2021; Mathur et al., 2019). Whenever specific options are preselected by default, before any user interaction, the **Pre-Selection** pattern is used (Gray et al., 2018; Mathur et al., 2019). Bösch et al. (2016) used this pattern with a different naming convention and labeled it **Bad Defaults**. This pattern poses severe consequences for the user by confirming practices often not in their interests (Di Geronimo et al., 2020). For instance, options for information sharing and email subscription in a digital service are selected by default. This pattern often comes in combination with hiding certain information about the preselected options in terms of conditions (Gray et al., 2018). Bösch et al. (2016) retrieved a pattern called **Hidden Legalese Stipulations**, which refers to using legal jargon to provide succinct but user-unfriendly terms and conditions. This pattern can be regarded as an extension of the **Hidden Information** pattern that disguises relevant information from a user (Gray et al., 2018). Using emotions in the digital context is also highly prevalent, as it can help to persuade users into specific actions and perhaps encourages them to change their initial intentions (**Toying with emotions**, Gray et al. (2018). The introductory example of Uber could be thus extended by labeling the logoff option of the pop-up screen as something like "No thanks, I want to lose $40 tonight" (Di Geronimo et al.,

2020). **False Hierarchy** refers to the visual or interactive antecedence of actions over others, such as options with text in a gray contrast that is mainly interpreted as not clickable. However, it is the opposite (Di Geronimo et al., 2020; Gray et al., 2018). Instead of using different visual effects, confusing language can also be used to trick users into making certain decisions (**Trick Questions**, Mathur et al. (2019). This approach comes mostly in hand with doubled-reversed or negative sentences to subtly invert the entire decision making-process (Özdemir, 2020). Mathur et al. (2019) listed one example of the opt-out process of email newsletters, where the text is similar to "We would like to send you emails. If you do not wish to be contacted via email, please ensure that the box is not checked." is displayed to the user. Advertisements could also induce customers into clicking on something that is not an ad (**Disguised Ad**). This pattern can be seen in practice on apps or websites where a click on any part of the service redirects to another entirely different website, making the first website an advertisement on its own (Gray et al., 2018). Mildner and Savino (2021) identified the **Guided Setting** instance in their study. For instance, any services that guide users through an inevitable process (e.g., adjustment of privacy settings) and list some options to them. But there is no guarantee that the user can see the full spectrum of available settings, as specific options (e.g., more privacy-friendly settings) might come at stake with the companies strategies (Mildner & Savino, 2021). The last instance in this Dark Pattern strategy is about making the users look dishonorable, ashamed, or stupid. Through **Confirmshaming**, particular actions will be framed with a damaging touch (often combined with the **Toying with Emotions** pattern), leading to an effective and powerful behavior change agent in the digital world (Mathur et al., 2019). Here, the authors mention an example when opting out of a newsletter yields a message like "No thanks, I hate saving money" to the user.

The following seven instances analyzed in this paragraph belong to the *Forced Action* class. They are any settings where users must perform specific undesired actions to continue using the digital platform. For example, numerous digital services require the user to recruit others and receive benefits or other incentives. This recruiting process can be linked to the accessibility of specific platform features, and this approach is subsumed in the **Social Pyramid** instance (Di Geronimo et al., 2020; Gray et al., 2018; Zagal et al., 2013). Taking a closer look at the privacy perspective, Bösch et al. (2016) found in their study that **Address Book Leeching** logically coincides with the **Social Pyramid** Dark Pattern. They considered the consequence of the pattern as an instance of its own, as data about users not being part of the network but is known to the network due to imported address books or solely mentions collected without their knowledge (**Shadow User Profiles**). Another prevailing practice among digital platforms is to use a telephone or mobile number in a Two-factor authentication (2FA). Usually, this technique primarily concerns security issues. However, services like Facebook use mobile phone

numbers also to place targeted ads for the user. In turn, the mobile number is used for a different purpose than the users might initially think. This instance is about **Privacy Zuckering** when users are tricked into sharing more information than they originally intended (Di Geronimo et al., 2020; Gray et al., 2018). Furthermore, several services in the digital world use **Gamification** - a way where access to certain service aspects can only be earned through the repeated and undesired use of the service (Di Geronimo et al., 2020; Gray et al., 2018). Zagal et al. (2013) mentioned the **Grinding** technique, where video game users have to repeatedly kill monsters to get crucial experience points to enhance the players' virtual character. Previously, specific skills were necessary to reach a higher level in games. However, to progress in the game, this technique is now a matter of time. This pattern can also be found in settings where users are subtly forced to buy particular skills to complete otherwise impossible game levels (Gunawan et al., 2021). This pattern is described as **Pay to Skip** by Zagal et al. (2013). Using the non-payment method requires users to spend even more time to achieve the same result they would get from paying the fee (Gray et al., 2018). This Dark Pattern instance can be seen in how games are delivered (**Pre-delivered Content**). Usually, the player purchases a game copy and should have access to all contents. This pattern becomes apparent when specific content is unavailable until the player pays an additional fee. Another technique found by Zagal et al. (2013) is called **Playing by Appointment** and refers to the amount of real-time defined by the game required to continue playing. For instance in farming simulations, where it takes some real-time in the game until some crops are ready to harvest. If the crop cannot be harvested within a game-defined time, the crop value deteriorates, so the player must play the game and not lose any crop yields. The **Pay to Win** (**Monetized Rivalries**) follows a similar logic to the **Pay to Skip** pattern. It exploits the users' competitiveness by forcing them to spend money to receive a higher rank on a leaderboard (Zagal et al., 2013). The last pattern in this corpus is named **Forced Registration** and can be found almost in every online service. Whenever a user wants to use specific functionalities of a service, it requires having an account. In particular contexts like social media networks, this is inevitable to use the service in a meaningful way. However, this procedure is unnecessary in most cases as the user serves as a valuable data provider for the companies (Bösch et al., 2016). Di Geronimo et al. (2020) further add instances to the list of the *Forced Action* class where users have to watch an advertisement to continue the service.

The *Social Proof* strategy indicates that people tend to evaluate their behavior based on peers. This Dark Pattern strategy includes three instances, namely **Activity Messages**, **Testimonials of Unknown Origin**, and **Impersonation**. The first pattern is of a transient, recurring, and attention-attracting nature in the form of a text message appearing on the top of a digital service indicating the particular activity of others on the website/app (Kim et al., 2021). Two types of messages can be further

distinguished. Static messages express that a person from location XY just bought a product. Dynamic messages can yield content that X amount of people just added the same item that the user is currently browsing to their shopping cart (Mathur et al., 2019). Digital services rely on testimonials to help customers to make their purchase decisions. However, the origin or the way how the testimonial was sourced and created is not transparent (Mathur et al., 2019). This Dark Pattern becomes even more apparent when the service provider does not offer any testimonial submission form or when the actual testimonials can also be found on other services. Zagal et al. (2013) further mention the **Impersonation** pattern and refer to situations where users get notifications about their friends' social activity. Still, they did not even perform this kind of displayed action.

The perceived value of a product and its desirability can be stipulated by using Dark Patterns of the *Scarcity* class, signal a limited availability or high demand for a product. Foremost, **Low-stock Messages** give hints about a products' limited quantity. This type of Dark Pattern can be even more sophisticated by hiding the exact amount. The results would be additional uncertainty and product desirability, triggering user impulse buying (Mathur et al., 2019). The aforementioned **Activity Messages** instance is mainly accompanied by the **High-demand Messages** pattern, which implies a high demand for the product and, consequently, the potential and instant out of stock if no purchase is undertaken immediately (Mathur et al., 2019).

The Dark Pattern strategy of *Urgency* collects all instances with deadlines imposed on sale actions to speed up the purchase decisions. **Countdown Timers** patterns dynamically display a deadline by counting down until the timer reaches its expiry date (Moen et al., 2018). However, in some instances, the exact and claimed to be expired offer is still valid even after the expiration. **Limited-time Message** patterns (e.g., a sale is supposed to end soon, but with no indication of the exact date and time) conclude the Dark Pattern instances list and seem similar to **Countdown Timers**. The significant difference lies in the static nature of these messages, as those do not come with accompanying deadlines (Mathur et al., 2019).

3.2.3 Dark Pattern Characteristics

After conceptualizing the different Dark Pattern strategies and the connected instances, it is relevant to identify the overlaps of different Dark Pattern types and what exactly makes them be considered "dark." Mathur et al. (2019) and Mathur et al. (2021) proposed six different characteristics that mark these evil design interfaces as "dark." It is important to note that each Dark Pattern type must not fulfill all six criteria to be considered a dark design interface. The first criterion concerns the **asymmetry** regarding all available user interface choices. This asymmetry can be seen in the unequal weights of a highlighted "accept-all" cookie button and, a

less contrast-emphasized reject button. The following criteria measure whether the effect of a user interface is hidden from users (**covert**). It answers whether the interface nudges users into making certain purchases without their knowledge. Moreover, a Dark Pattern type is considered "dark" when it meets the **deceptive** criterion. Whenever a user interface induces wrong beliefs with the help of affirmative misstatements or intentional omission of information, the Dark Pattern is considered **deceptive**. For example, Mathur et al. (2019) mention time-limited discounts, which remain in place after the timer expires, despite creating a sense of urgency among the users. Another facet of Dark Pattern characteristics is the **information hiding** intention behind evil design interfaces. The design interface is considered dark whenever a design setting obscures or delays crucial information that the users need for decision-making. Furthermore, whenever the users are **restricted** by the set of available choices in a user interface, the design interface can be regarded as a Dark Pattern. The last attribute is disadvantaging and treating a specific user group differently from another group (**disparate treatment**). One example might be seen in video game environments where users can pay to skip the payroll and advance in their game progress to gain advantages over users who did not pay or, even worse, cannot afford to pay the fee (Mathur et al., 2021).

Baroni et al. (2021) used a **social-technical** approach in their study to analyze Dark Patterns and conclude that the intertwined way is necessary to comprehend the phenomenon entirely. In the social world, Dark Patterns cause emotional and economic distress as it supplants the user value in favor of the shareholder. Furthermore, they delve into users' preferences to take advantage of them by, e.g., shaming them emotionally for not pressing a button in a digital service. The technical lens is more the structure for the users to understand and use a system. Both perspectives give a more unified vantage point on the domain of Dark Patterns. The technical perspective can further be expanded by analyzing the pervasiveness of Dark Patterns on different modalities. Nowadays, digital services are not solely available via a web application. Digital ecosystems and the revolution of smartphones also paved the way so that those are also usable with different modalities. Gunawan et al. (2021) used a **technical perspective** and analyzed whether the platform design matters regarding the prevalence of Dark Patterns. The authors observed that most Dark Patterns appeared more frequently in smartphone applications than in web browser.

3.2.4 Cognitive Biases

The main reason for the effectiveness of Dark Patterns resides in the ineffective human rational choice model itself. Due to bounded rationality, humans cannot consider all relevant information and transform these into an evidence-based decision (Waldman, 2020). This section associates the aspects mentioned earlier to get a

unified view of which cognitive biases are part of the underlying mechanism of the listed Dark Patterns. Although interestingly, the utilized cognitive bias differs among Dark Pattern strategies. There are only a few variations of the decisive classes. *Nagging* uses the **Framing Effect**, which relates to how information is presented to the user, either good or bad (Luguri & Strahilevitz, 2021; Moen et al., 2018). In the introductory example, Uber drivers are nudged into driving longer by framing the profit difference to an arbitrary sum with the help of the engine gauge.

The *Obstruction* strategy works because of the **Sunk Cost Fallacy** and the **Hyperbolic Discounting** cognitive biases (Mathur et al., 2019; Moen et al., 2018; Waldman, 2020). The *Roach Motel* and *Price comparison prevention* patterns utilize the **Sunk Cost Fallacy**. This bias refers to sunk costs that cannot be recovered once they have been incurred, as the user has already invested a certain amount of resources like emotions, money, effort, or time into it (Kim et al., 2021). The last two instances of this class, namely the *Intermediate Currency* and *Immortal Accounts* pattern, mainly exploit the **Hyperbolic Discounting** bias as these patterns are more about overvaluing current rewards and, contrarily, disproportionately discounting future costs (Waldman, 2020).

The *Sneaking* class is the strategy with the most variations among the used cognitive biases in the corpus of this thesis. The *Sneak Into Basket* and *Forced Continuity* patterns employ the **Default Effect**, as a preselection of options can be considered very efficient for nudging users due to inertia (Mathur et al., 2019; Moen et al., 2018). The *Hidden Cost* instance heavily relies on the **Sunk Cost Fallacy** bias. A user searching for products in digital services has already invested resources (e.g., time, cognition), so users tend to go ahead even with additional fees (Kim et al., 2021). The *Bait And Switch* pattern uses the **Framing Effect** instead by emphasizing the expected behavior of design elements, respectively (Gray et al., 2018).

Instances of the *Interface Interference* class mostly capitalize on the **Framing** and **Default Effect**. Using **Anchoring** is highly prevalent in *Aesthetic Manipulation* patterns, as the first salient information displayed to a user is crucial for future decisions (Mathur et al., 2019; Özdemir, 2020). *Preselection* patterns have certain privacy-invading or user-value supplanting options preselected by default using the user's desire to procrastinate to save time and effort (Berger et al., 2022). *Trick Question* patterns work similarly, as they first put default options in place that are combined with confusing language framing the user, thus, nudging them into particular behavior (Mathur et al., 2019).

Forced Action instances manipulate users by applying **Framing** and **Hyperbolic Discounting**. *Privacy Zuckering* Dark Patterns, like the 2FA of Facebook, as mentioned earlier, connote security by framing the possibility of having a secure account. Users, in turn, underestimate the future risks of disclosing this personal information used to place personalized ads (Waldman, 2020). Within this category of Dark Pat-

terns, **Framing** is used as a powerful vehicle to make the **Hyperbolic Discounting** cognitive bias work even better.

The *Scarcity* and *Urgency* class instances rely mainly on the **Scarcity bias** by making products appeal to have a superior value with the help of creating a potential fear of missing out (Mathur et al., 2019). The last strategy (**Social Proof**) in the corpus heavily capitalizes on the **Bandwagon Effect**, as users are likely to stick to actions that other users performed (Kim et al., 2021; Mathur et al., 2019).

3.2.5 Effectiveness

The question of how effective the Dark Pattern construct is in digital services regarding the usage of cognitive biases has been neglected so far. Luguri and Strahilevitz (2021) state that the apparent proliferation of evil design techniques in e-commerce is a piece of evidence to force users into things they might not otherwise do (**undesired compulsion**). They further concluded that relatively sophisticated Dark Patterns pose more threats, as they influence large quantities of customers to do things without provoking annoyance. Furthermore, the **legislation** seems to be behind the development in this area as well, as prosecution is costly and techniques are limited in effectiveness (Chromik et al., 2019). Personalization is also an issue as it allows even more malicious usage of Dark Patterns against users. Besides, as companies have to cut their costs and **cheap and effective sales techniques** can ease the corporate budget spent on acquisition and sales (Luguri & Strahilevitz, 2021).

3.2.6 Usage Motivation

Until now, the question of the main reason behind using these dark design interfaces in digital services has only been touched to a limited extent. Mathur et al. (2019) state that users in an online context decide differently based on how the same information is presented and specially framed. They further conclude that Dark Patterns can **trigger impulse buying**. Consequently, companies employing such deceptive design techniques can increase their profit. Chromik et al. (2019) advocate a similar stance. They express that Dark Patterns are used deliberately to delude and circumvent users to increase their corporate profits, but within the limits of the law (**legal profit increase**). Most usage motivation factors can be linked to direct or indirect **revenue acquisition**. In less common cases, nasty design techniques are used by some designers to **shock, disgust, or attack users** with their services (Conti & Sobiesk, 2010). They further mention selling a product/service, increasing brand recognition, gathering personal user information, and obfuscating legally mandatory data users hesitate to share as usage motivation. Luguri and Strahilevitz (2021) follow a similar notion, as they consider one of the leading motivational aspects of using Dark Patterns is the **increase of the companies' profit** by prompting users to

purchase products they did not intend for or by revealing more information about themselves. Speaking about increasing profits in the digital world, where free services are an inherent part of the market, excessive **collection of customer data** that can be accumulated and used for analyzing purposes remains the main form of revenue (Moen et al., 2018). The aggregated data can tailor individual digital services, and firms can sell targeted marketing. Due to the ubiquitous of apparently free online services, this data collecting and analysis part is the primary income source for many companies. Therefore, they heavily rely on users to share as much information about themselves as possible. However, users are not aware of how their data is used, but they are forced to share them with the help of Dark Patterns. The emerging dilemma can refer to the earlier mentioned information asymmetry on digital platforms, in which the power between the user and service provider is unevenly balanced in favor of the latter.

4 Semi-structured Interviews

This section is dedicated to answering the second research question about end-user accounts related to Dark Patterns and is organized as follows. The first passage briefly describes the underlying methodological approach to answer this research question. In the next step, the results of the method are collected, categorized into themes, and presented accordingly.

4.1 Methodology

The second research question intends to increase the awareness of the Dark Pattern phenomenon by exploring how end users perceive, experience, accept, and respond to Dark Patterns. The research approach is based on individual interviews and focus group discussions. In this way, it is possible to retrieve a thorough understanding of the participants' experiences and mindsets to provide a deeper understanding of this social phenomenon (Gill et al., 2008). The approach closely follows the work of Gill et al. (2008) and Newcomer et al. (2015) and intends to unravel insights regarding the awareness of practices and which Dark Patterns are considered acceptable or ethically problematic.

All participants were chosen subjectively but in consonance with having a homogeneous participant set (Newcomer et al., 2015). Furthermore, different control variables help to answer the second research question. These variables consist of the participant's age, technical affinity (following a Likert scale with low, medium, or high options), and usage frequency of digital services (using a Likert scale with rankings of never, occasionally, or frequently).

Typically, collecting qualitative information and the proceeding interpretation of the data is subject to certain biases. First, generalization from the results to a broader population is difficult, as a participant set usually does not adequately represent all users. Second, considering one background, e.g., academics with technical knowledge, creates another implicit bias in combination with a more narrowed down age range (Gill et al., 2008). Thus, three focus groups have been interviewed, each consisting of five people aged 22 to 56. Also, the selection of focus group members was made regarding their technical understanding. Group one consisted of five software developers and telecommunication engineers, and group two included less technically savvy people with different jobs, where IT is only end-user-centered. Group three combined people with a more general and broader technical understanding, e.g., people in executive positions. This separation in terms of technical understanding and age allows having a universal picture of the end-user perceptions from a different point of view. Also, it makes it possible for the participants to fully accommodate themselves and to get involved in the discussion to its full extent (Gill et al., 2008).

The participants of individual interviews were chosen by considering a balanced mix of technically savvy people and their counterparts with different backgrounds. Some have university knowledge, some are already retired or did apprenticeships and have working knowledge in various industries. The age range of the interviewees was between 18 and 66. This selection prevents certain biases and allows a broader view of end-user perceptions. Moreover, there was no overlap between the focus group participants and the individual interviews, meaning that different people participated in the interviews and focus group discussions. In total, 20 personal interviews were conducted, each lasting between 30 and 60 minutes, and three focus group discussions lasted 60-90 minutes. Regarding the used language, the researcher needed to use the tongue where the participants could fully express themselves and avoid any language-related misconceptions. For this reason, all focus group discussions and 11 out of 20 individual interviews were conducted in German. The rest was carried out in English whenever the participants could not speak German or it was not their mother tongue. This approach ensured the participants could fully express their thoughts and feelings towards the topic. The discussions of the focus groups took place face-to-face, whereas some individual interviews were conducted via Zoom.

All gathered data were rendered anonymously and handled confidentially to align with ethical research principles. Therefore, each participant was assured of ethical research principles of confidentiality and anonymity, to which they consented. With the participants' permission, all interviews were recorded and transcribed afterward. This approach protects against any biases and compiles a list of evidence about the sayings of the participants. The quotation of interview contents in the rest of the thesis will be aligned to the following scheme: Quotes from individual interviews will use the abbreviation of *Px*, where x stands for a number from one to 20 (participant

index). Excerpts of content from focus group discussions will use with the *FGx-y* pattern, where x represents a number from one to three (focus group index), and y is a number from one to five (participant index).

The data collection process used a flexible interview guide that briefly outlines the planned topics and questions, organized in the following six themes. The interview guide was used for multiple groups and interviews. Thus, it was tailored to each specific interview accordingly. Appendix F lists the complete interview guide.

1. Theme: Awareness, Harm and, Prevalence of Dark Patterns
2. Theme: Worries through Manipulation of Dark Patterns
3. Theme: Acceptance of Dark Patterns
4. Theme: Future Habits and Usage of Digital Services
5. Theme: Countermeasures for Dark Patterns
6. Theme: Company Perspective

For both the focus groups and individual interviews, the process started similarly with a short introduction welcoming words and ethical principles of the research, introducing all participants and thereby collecting the control variables like age, technical affinity, and usage of digital services. Before the interviews, all participants were invited with a rather generic topic description of *Manipulation in digital services*. Then the researcher asked some preliminary questions concerning experiences related to holiday bookings or online shopping to slowly advance the interview partners to the topic of Dark Patterns. Next, a short PowerPoint presentation illustrating three to four Dark Pattern examples was shown to the participants. Each screenshot was equipped with a fitting background story so that all participants could familiarize themselves sufficiently with the cases. Most participants were not familiar with working in gig economies. Therefore, the Dark Patterns were selected from simple everyday situations in digital services so that the participants could familiarize themselves better and avoid loose ends attached to unfamiliarity. As the interview guide was tailored to each conversation, the examples were adapted according to the interview language but used all the same underlying Dark Pattern strategies. Therefore, the samples consist of common Dark Patterns following the *Scarcity*, *Urgency*, and *Social Proof* strategies, more emotionally oriented instances, and more sophisticated examples where crucial information for a decision was hidden (Interface Interference). Appendix E lists all the used examples.

After showing the instances, the researcher gave a definition and a short recap of the *Dark Patterns'* impact on the examples presented earlier to the participants. In the next step, the interview questions were asked by recommendations of Gill et

al. (2008) in a neutral and timely manner, by not pushing the participants and the discussion in specific directions. After asking a question, the researcher remained passive to let the participants fully express their thoughts and feelings without any time pressure. The questions were asked by following a flexible scheme based on the course of the interviews and the six interview themes mentioned above. Besides, this sensitive topic about deception could quickly bring up any emotions hampering the participants' willingness to share their feelings and perceptions. To avoid this, the researcher carried out three pilot rounds. The retrieved feedback was incorporated into the interview guide before starting with the actual interviews.

The initial research design intended to analyze Dark Patterns and end-user perceptions in the gig economy context. However, as it was challenging to carry out the semi-structured interviews solely with gig workers, this thesis followed a different approach. Considering the demographic elements of gig workers (e.g., Uber taxi drivers), like mixed age, less to moderate technical affinity, and not necessarily academic backgrounds, the participant set included several people that match these characteristics (Lee et al., 2015). Also, the interviewees were exposed to Dark Pattern instances from their daily lives instead of showing examples common in online labor platforms like Uber. This approach mitigates familiarization problems for the participants, as they just had to imagine themselves in everyday life situations instead of projecting their thoughts into hard-to-grasp and unfamiliar working situations. Furthermore, it does not restrict the participant from expressing themselves.

4.2 Results

This section represents all the insights from the interviews. In terms of structure, the remainder of this chapter is organized by the different themes of the interview guide. In general, most participants use digital services at least occasionally to have a certain sense of the digital world.

4.2.1 Theme 1: Awareness, Harm and Prevalence of Dark Patterns

In terms of the general awareness and prevalence of this manipulation, the results draw a clear image, as most participants could recognize some manipulation. This visual manipulation felt very familiar for some interview partners, as they have "[...] seen this a thousand times before" (P2, 00:07:46), "[as] it is something that is kind of all over the Internet" (P13, 00:18:57) or "they are very familiar to me, I actually see them very often" (P15, 00:12:32). This pattern could be recognized among all age groups and without regard to technical affinity. Notably, older participants responded to the shown Dark Pattern examples with some caution (P4, 00:08:26) and relied on intuition (P6, 00:08:26). Some participants mentioned that they were not surprised

to experience manipulative attempts and defer this to their natural skepticism (e.g., P1 (00:07:23), P17 (00:12:16)).

A few people ascribed this kind of manipulation directly to marketing, as "[dark patterns] are usually very consciously chosen marketing strategies" (P17, 00:12:16). Or "[it is] all psychology, marketing [...] with marketing you always have to be careful" (P1, 00:10:38). Also, some participants draw a line to existing manipulation in non-digital domains: "I think it's everywhere, even in the grocery store" (P20, 00:19:51) or "that has always existed" (P6, 00:07:54). Furthermore, most participants tolerated this manipulation, as it was just visually perceived but not questioned (P3, 00:13:53). Others did not care about it beforehand (FG2-3, 00:22:14) or experienced it as usual (P14, 00:17:15).

Moreover, the control variable usage of digital services could not differentiate in this setting, as people with less or only occasional use of digital services could also identify some manipulation (e.g., P7, 00:13:13). However, most people were not familiar with the term Dark Pattern itself (e.g., P6 (00:07:15), P16 (00:15:05)). Only for a minority, mainly more technically savvy people, is the term commonplace (e.g., P19, 00:18:20).

Considering which type of Dark Patterns was perceived as more manipulative, there was no clear consensus among the participants. Some expressed that the hotel room example with *Scarcity & Urgency* Dark Pattern was "very familiar" (FG1-5, 00:21:36), associated directly with well-known brands such as booking.com and observed it as more pleasant than the others due to existing familiarity (e.g. FG1-3 (00:24:30), P13 (00:22:08), P20 (00:17:15)) or through appreciation of the design itself (e.g., P7 (00:10:16)). To continue this thought, some people indicated that even though the hotel example creates some pressure situation, the participants could handle it more because they have to decide before clicking and are more in the power of their decisions (e.g., P2 (00:08:42), P3 (00:10:54)). Contrarily, through the scarcity situation in the hotel example, some articulated the need of more information and input for their decision causes stress (e.g., P1, 00:12:38). Among the remaining Dark Pattern examples, the subscription sample was indicated to be more manipulative, as it uses colors in combination with feelings (P18, 00:16:06) and "makes you feel bad when you click on it" (P16, 00:13:09). Regarding the least recognized instance, the *Hidden Information* Dark Pattern, there was at least a consensus that this example was "more deceptive" (P10, 00:13:18) as it was "so hard to spot the other option" (P13, 00:22:08) or, speaking differently, as crucial information for a well-considered decision was missing.

The emotions felt by participants while being exposed to the Dark Pattern instances can be described as "negative feelings" (P3, 00:13:25), "the feeling of distrust" (P12, 00:13:13), "feeling annoyed" (FG2-2 (00:20:20), FG2-4 (00:23:31)), "[the] feeling of urgency and it was pressuring" (P13, 00:20:19). Others state that they experi-

enced these Dark Patterns as "intrusive" (P12, 00:13:34) or as a "betrayal" (P15, 00:13:52). Interestingly, the result from the technical focus group revealed a more prosaic and worldly-wise image of felt emotions (e.g., FG1-4, 00:25:05). Compared to the non-technical focus group, the overarching felt emotion was annoyance (e.g., FG2-2, 00:20:20), which is in line with similar results from less technically savvy people (e.g., P9, 00:12:09). Among the shown instances of Dark Pattern, notably in the *Scarcity-Urgency* Dark Pattern, the most participants felt being pushed (e.g., P7 (00:11:35), FG1-2 (00:13:07)) and have the desire to move fast, as they do not want to miss out on something (e.g., P18 (00:11:42)). However, there was also a more an objective perspective on the felt emotions, as one participant has acknowledged these examples as being "unprofessional" and "you are flooded with information and think to yourself, this could also go much faster if the facts were simply presented" (P16, 00:12:09). Some participants mentioned the reversal of their initial decisions, as they felt not wholly free to undertake their own decisions and could not choose for themselves (P18, 00:02:37) or the matter that their intention from the beginning "[...] is put in the background and something else is demanded by the other" (FG3-3, 00:26:25). Also, there was some resignation towards this manipulation, as "[...] you choose the simplest variant, because it means the least effort" (P9, 00:10:43).

The last question in this theme deals with the harm that Dark Patterns would cause the participants. Here, the results are mixed, without any variance in terms of technical understanding or age, as some participants mentioned financial aspects, as "[...] you might suffer financial damage (P2, 00:10:42), "I ended up spending more money" (P5, 00:15:20) or "I would spend money I would never have wanted to spend (P8, 00:12:59). Interestingly, only a few people mentioned that it is still their choice whether they spend their money or not (e.g., P13, 00:24:28). These Dark Patterns create some pressure and stress among some participants, or they are unpleasant and time-consuming (P6, 00:11:15). These aspects can be summarized as emotional damage (P11, 00:19:10). One participant put the sequence of caused damage also in order, as the manipulation does not cause "emotional damage because the financial damage comes first" (FG3-5, 00:30:51). Another participant mentioned the potential future harm, especially the harm "that my data is given away. That could lead to spam emails and messages from third party services" (P15, 00:18:05). Besides financial and emotional harm, one participant couldn't address real harm, as "you're kind of exposed to these threats frequently. And so you're just getting used to it" (P19, 00:21:11). The earlier found feeling of resignation is also visible in this aspect, as the non-technical focus group discussion brought up the thought of what harm it would cause for anybody, as they are not a person who is somehow a public figure or a politician (FG2-5, 00:24:46). The technical counterpart discussion disclosed a relevant aspect in the *Hidden Information* sample deliberately misleading participants, as people clicking on the disable button, will stay in the belief that

their data has been eradicated, but actually, the direct opposite is the reality (FG1-1, 00:31:46).

4.2.2 Theme 2: Worries through Manipulation of Dark Patterns

This theme covers aspects of how participants feel after realizing they have been tricked into something through Dark Patterns, that they did not intend to do, followed by the blame for the resulting fault. Generally speaking, some participants were not able to fully get behind the nature of the manipulation and spot the Dark Patterns in the first instance, after the explanatory part of the interview, this changed (e.g., P3 (00:12:16), P20 (00:18:23)).

The shared feelings of the participants can be described as being annoyed, mad, angry and disappointed in themselves (e.g., P3 (00:15:49), P5 (00:17:10), P7 (00:14:40), P11 (00:22:24), P14 (00:28:53)), feeling upset (P12, 00:17:21) and "a little bit ripped off and steerable" (P2, 00:11:43). Another person would find it as a pity if they bought something without intention (P1, 00:16:36) or could credit some "mini admiration like no other" to the service provider (FG3-1, 00:37:48). These results are homogeneous among the whole participant set without differentiating in terms of age, technical understanding and, usage frequency.

Participants with a low to medium technical understanding immediately presented some counter strategies, as they would try to cancel, return a product or change the pricing plan of a subscription (e.g., P9 (00:16:46), P15 (00:19:09), FG3-5 (00:35:03)). One participant expounded on feeling "probably surprised at first, but then I would start the investigation to find out what exactly that was" (P8, 00:14:14). Older participants also questioned the further usage of the service after experiencing manipulation through Dark Patterns (P6, 00:12:24) and have a lower tolerance, as they would "complain somewhere" (e.g., P6 (00:12:41)). A user with only occasional usage of digital services would immediately quit or try to reach out to the service provider directly, instead of relying on a chatbot (FG1-4, 00:32:55). There was a feeling of resignation noticeable, as one member of the technical focus group, would call reaching out to the service providers as "[...] a wasted effort and my money is gone in the worst case" (FG1-1, 00:34:02). Another more technical-oriented participant indicated that possible actions depend "[...] on how much and how big the impact is" (P13, 00:28:02), as Dark Patterns are "[...] so widely used that you are blind of it at some point" (P13, 00:33:03). In contrast to the non-technical focus group discussion, with particular emphasis on the hotel example, there was even some sort of acceptance of this manipulation, as long as there is a cancellation possibility for a longer time period (FG2-2, 00:26:02).

The second question in this theme reveals a three-folded image regarding who is to blame after recognizing the felt manipulation. The majority of technical-prone

interview partners blamed themselves, as the Dark Patterns "[...] are well done, but [they] didn't really look carefully enough" (FG1-3, 00:34:54) or "I am full of age and fully legally competent. If I made the decision that was my free decision" (P4, 00:13:28). One participant also mentioned a contradiction in these statements, as it is "actually strange, because I was manipulated, but I will still blame it on myself, actually I should blame the company instead" (P14, 00:29:14). Besides that, also technical less savvy participants would chide the users as "[the] person who ultimately clicks on confirm is always to blame. In that case, it would be my own fault" (P8, 00:14:54) and "[...] it's actually in your responsibility to look for the right button" (P15, 00:20:16). Mostly technical less and medium knowledgeable interviewees place a shared responsibility on both parties and consider the blame as a "[...] mixture between well done and self-inflicted" (FG2-4, 00:28:16) or blatantly as 50:50 (P17, 00:18:23). This can also depend on the impact in the end. "If it's really just colors or highlights being set, then I will blame myself because as long as I'm given the options to click on the right thing, it's my own responsibility [...]" (P16, 00:15:20). The aforementioned insight of resignation can be continued in this setting, as one person blames a 40:60 proportion but "can't really do anything about that" (P10, 00:17:55). Especially older and less technical experienced participants express a clear separation of accountability, as "[the service provider] added that, I read over it and clicked too fast" (P7, 00:16:13) or "partly it's your own fault in some situations if you don't question it so closely and click on it right away" (P9, 00:16:07).

The usage of age and education metrics could retrieve another compelling observation. The blame can also be seen as shared concerns, as "[every] person who grows up digitally, I would rather blame them, because they would have to know how things work and just be attentive. An older person, I would not blame now" (P2, 00:12:47). What is more, "[...] there is also some responsibility for the end-users to educate themselves, because this is so common everywhere" (P13, 00:33:03). To conclude with the last perspective of this spectrum, only a few participants with medium technical knowledge put the blame directly onto the service provider. They think that it was not their fault due to being "taken advantage of" (P3, 00:16:16) or being fooled despite thoroughly informing themselves beforehand (P11, 00:22:24).

To conclude this theme, there are also some rare opinions, as one interviewee credits the fault to the designers for designing it in a manipulative way. But in the same way, there is also blame on the person to be more aware of such manipulation due to previous experience with online services (P10, 00:17:21). Lastly, one participant responded to see this setting not in such a negative way, as the companies "are always trying to do the best they can to sell something. I assume that's not a bad thing at all" (P1, 00:14:02).

4.2.3 Theme 3: Acceptance of Dark Patterns

The current theme collects opinions about the possible acceptance of manipulation through Dark Patterns. The interviews yield versatile insights and offer a high variance related to technical affinity, usage frequency, and age control variables. In general, some manipulation is also known from the past (e.g., FG2-3, 00:37:35), and this prevailing manipulation attempts through Dark Patterns are regarded as acceptable, as long as it can be classified as advertising (e.g., P16, 00:17:48). To a large extent, by considering all technical levels and age ranges, the participants would agree on these sort of manipulation through Dark Patterns, forasmuch as there is a personal advantage prevailing (e.g., P4 (00:17:44), P19 (00:26:21), FG2-2 (00:36:01)), something is cheaper (P20, 00:22:38) or useful (P18, 00:27:44).

Furthermore, peculiarly the hotel example was considered tolerable, as most participants have seen it many times, with the result of not having any impact on them anymore (e.g., P1 (00:19:37), P2 (00:14:16)). One participant mentioned that these Dark Patterns can be acceptable. However, it requires much work in terms of transparency (P13, 00:39:12). Moreover, two interviewees expressed their complacence for cookies as a helpful tool to display ads with relevant products that help them to find the right product with less effort (e.g., P11 (00:25:27), FG2-4 (00:31:54)).

In the interviews, the participants were also asked about their impressions regarding a standard digital market feature that shows them what other customers also bought. This functionality is a more subtle form of *Social Proof* Dark Patterns and can trigger additional purchases but could also yield interesting input for the users. Here the results are mixed, as significantly less technically savvy people consider this as valuable, as it helps to make the right purchases in unfamiliar domains and to prevent overlooking certain aspects (FG3-4, 00:42:43). Also, more technical-experienced interviewees considered this feature as "quite workable" (FG1-5, 00:39:45). Furthermore, it can even help to save shipping fees when combining purchases (P13, 00:37:49). However, there was a consensus among these participants that it remains to be their own decision if they buy the additional product or not (e.g., P4 (00:17:44), P9 (00:18:40)). Besides, there is also some rejection towards this feature, as some considered it as "[...] no useful information. If I'm looking for something, I look for it specifically" (P8, 00:16:31) or simply ignorance about what other people bought (e.g., P11, 00:26:23). To conclude this spectrum, some participants also remained neutral concerning the applicability of this feature for themselves (e.g., P7 (00:18:08), P19 (00:25:49)). In a broader sense, trust is inevitable for transactions between the user and the seller (P6, 00:13:55) - if this trust is shattered, users would consider that an "absolute no-go" (P2, 00:14:16).

This theme also included a sophisticated question to reveal ethical implications of the users about their habits when using a free online service that collects and sells

their data. This assessment conceded some hypocritical stances among parts of the participants. Significantly less and medium technical savvy people would reject using these sorts of services, despite being signed up for services like social media employing these practices (e.g., P3 (00:18:46), P5 (00:21:14), P12 (00:21:23)). Likewise, if people already knew beforehand that specific manipulation practices occur on services, they would not use the service (e.g., P16 (00:19:14), FG3-2 (00:44:47)). The focus group discussion in the mixed-group exposed some convenience aspects, as there is some tolerance regarding these practices because it is often accompanied by convenience (FG3-4, 00:40:44) and usually "you turn a blind eye to it, because it means effort" (FG3-4, 00:44:55).

Another overall notion coincides with the initial thought of resignation regarding this ethical dilemma. Participants with all technical depth levels state that "as soon as I use any internet platform, you have to accept it" (P9, 00:17:26) and "it's actually how the market works [...], and they are free to use the tools they want to (P5, 00:19:22). The resignation spreads even more, as "there is not really an alternative" (P14, 00:32:27), "[...] you have the feeling that you can no longer get out of the vortex" (P2, 00:17:08) and they cannot do anything about it (e.g., P20 (00:24:27)). Besides the feeling of despair, some participants also approached some active questioning of the situation. One common thought was the desire to "[...] know what data is actually processed" (FG3-3, 00:45:24) and to whom the data is sold (e.g., FG2-2 (00:37:40), P7 (00:19:13)). Others state that despite knowing the selling practices of services, they either hope to avoid this by paying more money for a product (P14, 00:32:27) or by relying on existing regulation (P18, 00:28:49). Notwithstanding being "super unethical" (P11, 00:27:49), some medium technical savvy interviewees thought "that most people don't really feel the harm that can happen to you when giving away those data" (P15, 00:25:41) and people are not aware of it when consuming these services (P17, 00:25:17).

To conclude this theme, some respondents with medium and high technical skills and of mixed ages were also reluctant to accept these practices by actively scrutinizing the nature of a free online services, and their income sources (e.g., P6 (00:17:25), FG1-2 (00:44:10)).

4.2.4 Theme 4: Future Habits and Usage of Digital Services

After experiencing Dark Patterns, it raises the question of whether participants still use the affected services and when the tipping point for them to quit is reached. This topic handles these issues. Across all technical levels, some participants decide to leave a service depending on the service's necessity. So speaking differently, if they need a product/service, they will continue to use it despite the experienced manipulation (e.g., P3 (00:19:55), P7 (00:20:01)). Otherwise, the participants would either

"stop using it" (P5, 00:23:00) or "try to find alternatives somehow" (P3, 00:22:14). The last approach is supported by many participants, especially those with lower and medium technical comprehension, and they consider it as a practicable approach to avoid the manipulation through Dark Patterns (e.g., P11 (00:31:11), P12 (00:25:19), FG3-4 (00:48:49)). However, the alternative service has to "be as efficient, and at the same time they respect these kinds of personal information and ethics stuff" (P11, 00:28:36). For other participants, the service offer will induce the future usage and "whether it gives me a good feeling" (FG3-1, 00:56:58). In the technical focus group, the prospective use of digital services employing Dark Patterns also depends on the fact, if the company had already applied Dark Patterns before the user started to use the service or if the firm is now heading in this direction (FG1-2, 00:46:35). But also laziness can influence the decision to remain using a service, "no matter how many scandals there have been" (FG1-1, 00:49:15). The preceding theme raised the thought of despair, which can be noticed among some technical-mixed skilled participants as well. They are gloomed about stopping using a service with Dark Patterns, as it is too late for certain services (P2, 00:21:25) or "you don't really have a choice in not using them" (P10, 00:25:32). Interestingly, more technically savvy interviewees brought up the thought of group behavior that is inhibiting their decision to quit a service. This behavior might be due to the impression that "a decision on your own has to be a mass decision" (P13, 00:41:08), and people continue to use services despite feeling manipulated due to group behavior (e.g., FG1-5 (00:39:45), P14 (00:34:53)).

The control variable of technical affinity differentiated the results regarding the potential tipping point for participants to leave a service. For less skilled interviewees, data selling (e.g., P3 (00:21:49), P20 (00:27:59), FG2-1 (00:45:11)), and financial harm (e.g., P10 (00:24:05), FG2-1 (00:46:03)) are considered to be unacceptable. Contrarily, the technical focus group is concerned with the absence of trust as a crucial indicator to turn one's back on the service. For instance, the collected and processed data is in no eminent relationship with the offered service (e.g., P19 (00:28:02), FG1-1 (00:48:49)) and the provider is not reputable (P4, 00:24:01). Some participants would rather stop during the registration for a service when they spot a possible overuse of their data (FG1-1, 00:51:40). Furthermore, if a company is performing leveraged buyouts and the user do not coincide with the value set of this firm, it would also mark a turning point (FG1-2, 00:48:40). But in general, the participants would leave a service when it "becomes too cumbersome" (P16, 00:20:04), "too penetrating and too intrusive" (P17, 00:27:54), when they are "tricked by some fine print into falling into a trap" (P12, 00:23:49) or when something very concrete, like a "blatant invasion of privacy" (FG3-2, 00:51:45) is affecting their real life (P13, 00:42:45).

To conclude this theme, all participants were asked about their stance regarding services collecting and processing sensitive data like bank accounts or biometric data. The results can be summarized as this approach would mean to cross a red line

and is out of the question (e.g., P6 (00:19:56), P7 (00:23:14), P16 (00:21:14), FG1-4 (00:50:15)). Bank data can be considered as "special personal data" (P9, 00:22:23). Therefore, "you expect a trusting and professional relationship from the bank" (P2, 00:22:18) and put "[in] trust that sensitive information is being taken care of" (P5, 00:23:50). Furthermore, if governments are behind this data collection and processing, some participants would also avoid using the service (P1, 00:25:16).

4.2.5 Theme 5: Countermeasures for Dark Patterns

The data represented in this chapter deal with whether it is possible to avoid manipulation through Dark Patterns and feasible remedies. Regarding the first statement, the prevailing echo of the participants across all technological skill ranks and ages was that it is hardly achievable (P19, 00:29:41), "difficult" (P2, 00:23:31), and due to the disseminated usage of Dark Patterns, they cannot skip them (P20, 00:27:59). Some medium technical skilled interviewees credited this form of manipulation to be a "condition of the market" (P5, 00:25:54). In contrast, the current legislation would hamper ways to tackle this kind of user control (P10, 00:26:43). Interestingly, older participants would recognize a way out of the situation but see it as the provider's turn. Conversely, they would not change anything because they would have to forego many profits (e.g., P7 (00:24:26)). Another person is also proposing that it is possible to limit this kind of manipulation. However, the profit thought cannot be eliminated (P3, 00:26:23). The technical focus group questioned the necessity of Dark Patterns when a "fictional honest company [is] just trying to sell their product and [you] think it's a good product" (FG1-2, 00:52:56). This discussion also revealed some tentativeness regarding the avoidance of Dark Patterns as it is challenging to draw the boundaries where Dark Patterns start and end (FG1-5, 00:53:55). The mixed discussion group came up with a unique thought about the potentially biased worldview of the younger generation, as they might struggle to perceive this as manipulation, instead of accepting it as normal behavior (FG3-4, 01:00:13).

Concerning possible countermeasures for Dark Patterns, many participants would see cognitive measures in the first instance. That means "being vigilant" (FG1-3, 00:57:29), being more careful, and reading things more carefully (P10, 00:27:37), "informing myself in advance" (P12, 00:26:18), "not just blindly clicking through [...] and always actively looking for alternatives" (P2, 00:25:00). Another idea would be to limit the time and amount of different services (P13, 00:45:54). Notably, older people rely to a large extent on their intuition (P6, 00:20:54) or gained experience (P4, 00:26:02). Besides that, only one older interviewee voted directly for any legislation changes (P6, 00:20:54), whereas especially less and medium-skilled people indicated the need for more transparency of how their data is collected and spread out (e.g., P18 (00:33:39), FG2-4 (00:55:22)). This transparency can be established via legislation

or education starting early in school. Another option would be the media (e.g., P3 (00:27:31), P17 (00:31:50)), maybe in the form of a mock exam (P14, 00:37:24). Furthermore, technical savvy interviewees came up with more technically oriented solutions like using the browsers' incognito mode (e.g., P1, 00:28:17), Virtual Private Network (VPN) (P1, 00:30:28), ad-blockers (FG1-4, 00:57:22) and "using several sites at the same time" (P2, 00:23:31).

At this point, the participants were asked during the interview about their stance regarding the usefulness of a browser feature that rates websites based on their usage of Dark Patterns. Here, the results are diverse, as less technical versed people admitted the versatility of such a tool, which could help them to be put "in an alert state of mind when you entered a website and you know what to expect" (P15, 00:31:15). This tool should be available already within the browser and not in the form of an extension that has to be installed again, as it would be "technically more complex for the normal user" (P6, 00:23:46). However, some older people brought up a painful subject in this context, as the vendors of browsers are also keen on earning money somehow (P4, 00:26:29). This would raise concerns of being manipulated by a tool that should prevent manipulation (FG3-5, 01:04:05). Other, more drastic solutions for dodging these behavior-influencing instances would be to stop using the service completely (e.g., P13, 00:45:04)

4.2.6 Theme 6: Company Perspective

To conclude the interview, the interviewer performed a perspective switch to the company's side. The participants were asked about the primary motivation, differences in company sizes and industries, and the responsibility within a firm to employ such deceptive design techniques. In terms of the main motivation, across all technical experiences and ages, the results showed with an overwhelming majority that profit motives serve as the main impetus for companies (e.g., P5 (00:28:58), P11 (00:36:04), FG1-1 (00:59:11)). Some interviewees also mentioned secondary intents that relate to profits, like generating "click-rates" (P4, 00:28:06), collecting data (P8, 00:23:15), "entic[ing] customers to long-term relationships" (P16, 00:24:59), "identify[ing] user behavior or generate leads for others" (FG3-4, 01:07:13). Collecting vast amounts of data also provides companies some power to have clues about "where the society is going" (P1, 00:31:25). These insights are in line with previous reported motivational factors listed in section 3.2.6.

Speaking about the aspects of company size and industry, where Dark Patterns are more prevailing, technical savvy participants expressed that bigger companies tend to use this form of manipulation more, as they have more resources and knowledge available than smaller ones (e.g. P1 (00:34:26), P4 (00:28:52)). Compared to smaller companies, the bigger players use Dark Patterns in a more subtle form (P2, 00:29:02),

have "smoother techniques" (P11, 00:36:33) or "might be able to get more advantage out of [the data]" (P19, 00:31:57). Contrarily, less technical versed interviewees saw more an equal distribution, as due to the lack of proper resources, smaller companies might then hire external people to implement these design interfaces (P6, 00:24:55) and they have to compete with the bigger ones (P3, 00:34:56). Concerning the industries, differences in technical affinity, and age faded away, as most interviewees stated that Dark Patterns are more widespread in industries with a more technological focus (P16, 00:27:46), the travel industry (P2, 00:29:02), and e-Commerce in general (P3, 00:36:23).

Lastly, the question about primary responsibility provides compelling outcomes. Technically less savvy people would put the blame directly on the marketing department (e.g., P10 (00:31:43), P12 (00:30:23)) or, generally speaking, "people who came up with the strategy" (P20, 00:31:21) or a "higher instance, a pretty highly psychologically trained instance" (FG2-1, 01:01:25). Moreover, they consider the software developer to be just "the executive power" (FG3-1, 01:11:48). Contrary to this, technically more advanced interviewees indicated "that it is very hard to point the finger at like a single person who's responsible" (P13, 00:50:48). They see it instead as a shared responsibility (P12, 00:30:59) and an "exchange of ideas" (P19, 00:34:17). More responsibility is credited to the strategic part of the decision to implement these deceptive techniques (e.g., P1 (00:31:29), P3 (00:37:03)). This kind of proportion could also change if the software developer is still creatively involved (e.g., P3 (00:37:26), P16 (00:28:37)).

5 Discussion

From the review above, several key findings emerge: In general, most participants recognized some manipulation, especially more common Dark Pattern instances from the travel industry, which were very familiar to all participants. Some participants could not perceive the Dark Patterns right away. Some could recognize the potential presence of Dark Patterns in the context after the interviewer pointed it out. The implication of this finding is also discussed in the work of Di Geronimo et al. (2020), who labeled the phenomenon as *Dark Pattern Blindness*. According to the control variable of technical affinity, most of these interviewees are less or moderately technically skilled, which implies that technical understanding can help to recognize and identify Dark Patterns quickly. As mentioned earlier, most interviewees observed manipulative attempts. However, the interview partners mostly did not know the term Dark Pattern. This finding coincides with previous research (e.g., Gray et al. (2021) and Maier and Harr (2020) and allows the researcher to retrieve an increased awareness of these patterns among users.

Nonetheless, technically less and moderately skilled participants rarely questioned the nature of Dark Patterns. Instead, they tolerated it. Given this background, the control variable of technical affinity could differentiate the results. One reason for this could be a different mental model that technically savvy people have compared to their less-skilled counterparts. Coupled with a higher cognitive activation, this mental model can help them question the nature of specific deceptive techniques. These results show that educational measures are needed to sharpen the ability to recognize manipulation. In addition, the participants' experiences in the digital world greatly influenced their perceptions.

Furthermore, the results also illustrate a valuable perspective on which Dark Pattern instances were considered more deceptive than others. In general, as long as the interviewees remained in the power of their decision, they accepted the particular Dark Pattern. However, the case of missing crucial information for a well-considered decision marked a turning point.

This analysis also revealed that notably older people (older than 40), even with less equipped technical understanding, relied heavily on intuition and life experience. A similar pattern was not obtained in previous studies. This is interesting, as some might expect these deceptive design techniques to trap older people with less technical understanding fully. Reasons why this scenario occurred could be a somewhat cautious attitude of more senior people regarding digital techniques and products, leading to certain reservations, which in turn could prevent them from being more affected. This kind of behavior was also observable during the interviews. However, in this context, older people could sharpen their awareness of different types of Dark Patterns to identify other online manipulation more easily.

It is also worth discussing the results concerning the general feelings of Dark Patterns. More technically skilled interviewees reacted less emotionally than less moderately versed people. This circumstance could be due to the existing familiarity because Dark Patterns are no longer out of the ordinary. Another explanation would be that technically skilled participants could have problems fully articulating their feelings. Regardless, this notion was noticed across all discussions with technically more competent people. For less and moderately tech-savvy individuals in the interviews, the most frequently mentioned feeling was annoyance, which ties nicely with previous findings by Maier and Harr (2020). Thus, the argumentation of this study, fostered by the results, is that people react angrily in the first instance, and this feeling may decline over time as people get used to certain types of Dark Pattern. Eventually, the feeling will subside, so these practices become normalized and part of the expected interaction flow when using particular services. Based on this stance, the interviews revealed some resignation, especially among less technical people, which can also be identified across other themes. A sparking question arises here, whether

these people have an insufficient awareness regarding specific practices in the digital domain.

The control variables of age and technical affinity used in the present study partly confirmed prior findings about worries through manipulation. Older people questioned the future usage of services where they experienced manipulation through Dark Patterns. This result highlights new insights about different age groups experiencing manipulation through Dark Patterns. The root cause for this behavior could be a particular value set that especially older people developed within their lives, which will be convulsed through the manipulation, leading to such an averting of the service. Less technical-versed interview partners immediately responded with counter strategies during the interview. This behavior might be explainable, as these participants might be vulnerable to this manipulation, and after getting caught red-handed, they tried to reduce the damage.

Attributing blame for the manipulation experiences provides more detailed insights compared to previous studies. Some participants place the blame exclusively on the service provider. This pattern is consistent with what has been found by Maier and Harr (2020). However, this report goes beyond previous work. The pattern, as mentioned earlier, was only ascertainable for medium-skilled and younger participants. In turn, older interview partners and more technically versed people saw a more shared blame responsibility. The first aspect could again be explained by adult thinking and life experience. In contrast, the second notion may be related to a deeper understanding of the essence of the nature of Dark Patterns, which is possible through more technical knowledge. Hence, more technically versed people acknowledge that being critical when interacting online and paying attention are crucial steps to evade these Dark Patterns. However, it is worth pointing out that there was also an ethically debatable contradiction in the blame, as participants blamed themselves instead of the service provider.

Regarding the acceptance of Dark Patterns, the results answer this question more on a generic level, as the experienced persuasion is subjective. Nevertheless, the report shows that the interviewees tend to accept this manipulation as long as it can be classified as advertising or as long as the Dark Pattern is completely visible. Moreover, the participants accept this manipulation when users have a choice or cannot decide for themselves due to a lack of information. Furthermore, regarding the hotel example, there was the fear of missing out on a tempting offer. Consistent with previous studies by Maier and Harr (2020), some manipulation is acceptable, as long as the personal advantage is considered higher. Here, the connection to the *Privacy Paradox* can be drawn, a phenomenon when people claim to care about their sensitive and personal information, but at the same time, they share them excessively to get certain benefits in return (Barnes, 2006; Kokolakis, 2017). During the interviews, most participants claimed to be highly sensitive with their data, but all of them stated to use social

media. In this situation, the *Privacy Paradox* fits in smoothly. One reason for this could be ignorance of certain practices in the digital world. This claim is supported by the results of the interviews, as participants wished for more transparency related to companies' collection and selling techniques. Again, this is consistent with previous literature (e.g., Kokolakis (2017)). Another explanation could be the interviewees' willingness to share sensitive data when the perceived risks are minor compared to the expected benefits (Draper, 2017). The results of this study also back up this claim, as the interview partners perceive the manipulation as acceptable, as long the personal advantages outweigh the derived risks. Lastly, this paradox may be further interpreted in the way of the reported resignation and inevitable confrontation of the manipulation of digital services (Draper, 2017). However, this does not exclude a general stance of rejection when such malicious techniques as Dark Patterns are in use.

Moreover, the present results highlight that technical-savvy people tend to analyze and scrutinize the digital sphere with more caution because they question the free nature of an online service and its potential unclear data processing strategies. They also consider whether a service they use has just introduced Dark Patterns or has already utilized them since their initial usage. Contrarily, the behavior of less-savvy interviewees resulted in the aforementioned *Privacy Paradox*, and the reasons to quit a service were more related to data selling scandals or financial harm. Some participants respond more when something undesirable happens to them instead of acting proactively. For technical affine interviewees, the absence of trust, chiefly using data that does not fit the service relation, was considered the central tipping point to stop using a service. If this trust is disturbed somehow, interviewees with a more technical background would immediately stop signing up for a service.

Considering the future usage of services, the interviewees tend to continue using a particular service if they perceive it as necessary. Prior to that, however, primarily less technical knowledgeable participants would try to find alternatives. In this setting, the results demonstrated an exciting finding, as the participants were not reticent in expounding their reasons to continue using a service, despite being tricked off. This insight also coincides with previous studies, such as Kahneman (2011), indicating that despite knowing their biases, people struggle to act against them. In this context, it is worth discussing the relevant fact about network effects hampering users in their decision to quit using a service. Some respondents indicated the need to continue using a service to stay in touch with their friends, despite knowing about the manipulative attempts they face while using their service. Staying connected with their friends and community is very important for many interviewees, even though it can be considered a trade-off with the people's perceived values from the services and their own value set. To put it differently, with the social influences and overall network effects, participants struggle to leave a service and supplant their values partly in favor of the perceived

benefits of using certain services. For users, these digital platforms deliver value and benefits, but they also reduce the users' remaining options and flexibility. This user dependence on services allows the companies to experiment with deceptive techniques without fearing user out-flux.

Based on the study results, the companies that pursue these manipulative strategies rely on one overarching motivational factor, profit. However, with the ubiquity of social media nowadays, this approach is risky, as customers' trust and credibility are likely to be reduced, based on previous findings. Regarding accountability for deploying such deceptive design interfaces, participants with a more technical focus reported a shared responsibility. In contrast, less technically-savvy people burden the executive departments such as marketing. This insight coincides with previous statements about the differences related to technical affinity among the participants.

There was a clear image regarding the company sizes, as leading firms use it more due to more existing and available resources. There might also be more knowledge about possible ways to monetize the data, which is not the case for smaller companies. Those would hire external contractors to use these techniques to compete with the more prominent players. However, having more resources and possibilities to use the data could provide certain big players also with the advantage of forecasting in which direction and how the society is developing. This issue raises several ethical concerns, which are presented at the end of this section.

The present study reveals insights into the potential avoidance of services using Dark Patterns and possible remedies. Regarding the first aspect, it is impossible to avoid and elude this manipulation, according to the participants. This pattern was also revealed by previous work by Bongard-Blanchy et al. (2021), Di Geronimo et al. (2020), and Maier and Harr (2020). There might be a variety of reasons for this. Several participants pointed out that it may be related to market structure with overall network effects or the widespread usage of Dark Patterns, especially in the e-commerce sector. For example, companies with larger market shares, on the one hand, have the resources to come up with newer and even more sophisticated design techniques. On the other hand, due to existing network effects, these companies can act as one may see fit and experiment with deceiving design techniques while not being afraid of users leaving. To overcome this, some participants mentioned that changes in the capitalist system would be necessary, requiring a strenuous effort across multiple industries and companies stepping down from their profit orientation. Investigating whether this approach would be feasible is out of the scope of this thesis.

Interestingly, some participants had difficulty thinking about possible countermeasures, as it was challenging to assess where Dark Patterns start and where they end. More precise separations of which Dark Pattern instances are more deceptive than others could help to address this question. Several participants across all technical levels proposed transparency as one key element to avoiding the manipulation

of Dark Patterns. Transparency about how data is being sold and processed and the influence of their collected information on their displayed offers would help them to understand certain functions in the digital world better. Legislation can be one step toward reaching this and should be built upon previous attempts to refine and sensitize people regarding processing their data. This thought is consistent with what was found in an earlier work by Mathur et al. (2019).

However, based on the results of the present study, this kind of manipulation literacy should be considered just as a starting point for countermeasures against Dark Patterns. Media coverage can help to achieve a broader awareness regarding this topic. People can use this as a platform to openly discuss the ethical usage of design techniques. Furthermore, publicly condemning and blaming businesses using these design practices can also be a way out of this situation. Speaking the language of risk can remedy the overly narrow focus of the discussion. Together with invoking ethical principles of companies, this could spark a turning point for companies to avoid any backlashes from potential scandals. However, the plain awareness of the existence of evil design interfaces is not a universal weapon against being deceived. Raising awareness in public might not be sufficient. Hence, businesses using unethical design techniques should be condemned from different angles. Early education starting in school should help create a framework that allows younger people to classify this manipulation more quickly and be better aware and prepared. Ethical design principles and implications of cognitive biases should be part of the curriculum for Information Science students, especially UX designers. For older and non-academic people, a focal point where they can report such Dark Pattern instances and share them with other human beings would be worth considering. Apart from that, the integration into everyday life is crucial, especially concerning the younger generation, as they would have otherwise a biased worldview and perceive this manipulation as usual.

The previous discussion unfolded several ethically problematic issues that need to be discussed. First, the interviews' analysis revealed a peculiar blaming placement after being tricked into something through Dark Patterns. The participants blamed themselves more instead of the companies. From a Kantian perspective, the companies violate the universal law by cheating the users through Dark Patterns. In other words, the companies trick users with Dark Patterns into unwanted actions and, therefore, break the underlying laws of the Kantian theory as cheating is not allowed (Kant, 1870). Consequently, the usage of Dark Patterns may be considered unethical, and people may be expected to blame the company instead. However, this reversal might be explainable, as the users of digital platforms fall for the companies' tactics and then blame themselves more than the companies.

Furthermore, the intention of using Dark Patterns is ethically equivocal from a company's perspective. According to the interview results, their primary intention of using Dark Patterns is to fool users to increase corporate profits. However, this intention is ethically fraught with problems, as it violates the categorical imperatives of Kant. According to this theory, companies should not use Dark Patterns to cheat on users to increase their revenues, as they deviate from the moral laws leading to unethical behavior. Furthermore, by fooling the users, companies do not treat the users as valuable human beings, portending another aberration from Kant's universal law. Lastly, the deliberate usage of Dark Patterns is also at stake with the third categorical imperative (Kant, 1870). Moreover, the fooling does not benefit the users. Instead, it is in favor of unearned benefits for the companies. Therefore, the intention of companies to use Dark Patterns is even more problematic from an ethical perspective. The analysis further revealed that manipulation through Dark Patterns is widespread, and firms use it deliberately to trick users into particular actions. This point of view allows marking the companies' behavior as clearly morally wrong. Kimppa (2020) steps further and classifies it more devastatingly as evil. Speaking differently, a lousy intent of the companies leads to unjustifiable systems, and the designers can be classified as evil without looking at the consequences. Conversely, accidentally good consequences of an evil intention do not justify the goodness of a system. Furthermore, as indicated earlier, intentions are difficult to assess ahead of time. However, in terms of Dark Patterns, it is different, as the results clearly showed that Dark Patterns are used to fool users, thus making them immoral.

Examining the software developer's perspective yields considerable insights. According to deontological theory, software developers implementing deceptive design interfaces would consider the removal of Dark Patterns from digital services a violation of their professional code of ethics (Kant, 1870). Despite knowing that users are being tricked into undesired actions by created Dark Patterns, the removal would be a form of lying or cheating, violating the universal law, and condemning the action as ethically unacceptable. In this setting, the weakness of the Kantian theory becomes apparent.

By looking from a consequence-based perspective, using Dark Patterns is also ethically questionable, as the consequences of Dark Patterns differ from the initial needs and expectations of the users. In the long term, the consequences of using Dark Patterns would imply that more unnecessary things are either being sold to or used by the people in society.

The results mentioned earlier also disclosed the privacy paradox, which may not be relevant from an ethical perspective. The circumstance of sacrificing something that people should not do does not pave the way for acceptance. Notably, as the interview results reflect, most people are unaware of the risk they are taking.

Besides that, forecasting possible development avenues for society through data collection can be considered ethically acceptable according to utilitarian theory, as it could produce the greatest good for the most significant number of people. Moreover, the collected data can help to improve technologies that are being used to enhance the everyday life of people. However, in the future, it is difficult or almost impossible to predict whether the consequences of the collected data through Dark Patterns will be good or bad (Stahl et al., 2014).

The question of potential parallels from the findings to end-user perceptions in the gig economy has not been discussed yet. Based on the previous in-depth analysis, it is crucial to point out that the underlying Dark Pattern strategies are the same for Dark Pattern instances in the gig economy and the shown examples. What matters is the final implementation of the design interface. However, the exploitation of cognitive biases remains the same. In this way, the retrieved results allow the transfer of potential end-user accounts of Dark Patterns in the gig economy. However, the context aspect of gig workers can only be alleviated to a minor extent. That means gig workers are using online labor platforms to make their living and are dependent on them. Therefore, exposure to Dark Patterns might trigger different emotions and results compared to everyday situations like the hotel booking example. This pattern is also in line with the work of Lee et al. (2015). Nevertheless, the results above can be generalized across different economies more abstractly, leaving space for future research agendas.

At this point, it is essential to note that this thesis should be interpreted with some limitations, which could be addressed by future research. First, the study could use only the current snapshot of the existing literature for the structured literature review. Therefore, one limitation of the present study includes that Dark Pattern instances, which have been created and identified afterward, are not covered by this thesis. Second, the participant chosen for the interviews were not related to the gig economy, which could lead to certain biases while generalizing the results. A potential future research avenue could be to draw on this study and extend the analyzed end-user accounts with a direct approach to the gig economy instead of using the indirect way. Third, the participants approached the screenshots with Dark Patterns during the interviews from a relatively objective perspective. Apparent limitations of this undertaking are difficulties in grasping the content to the full extent and the hindering of interviewees from expressing themselves perfectly. Future research could try to establish a more practical setting to overcome these issues. For instance, users can be directly exposed to Dark Patterns while navigating and performing specific tasks on particular digital services. An additional aspect to consider is the interview language of English, which was not the mother tongue of a few participants. The feasibility of

answering the interview questions in a second language restricted some interviewees from fully sharing their thoughts and opinions. Even though these participants are completing their education in English, conducting a similar study may be helpful where they can use their first language to express their feelings about this topic. Lastly, the study used a qualitative approach to respond to calls from the research community on how end-users perceive Dark Patterns. An appropriate next research avenue would be to conduct a quantitative study to determine what types of users are more susceptible to manipulation or if there is a varying degree of reaction from different user groups. Finally, future work may focus on potential legislation issues to address the unmet desire for transparency regarding Dark Patterns and data usage in digital services.

References

Abdelkafi, N., & Täuscher, K. (2016). Business models for sustainability from a system dynamics perspective. *Organization & Environment*, *29*(1), 74–96.

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines: The simple economics of artificial intelligence*. Harvard Business Press.

Arkes, H. R., & Ayton, P. (1999). The sunk cost and concorde effects: Are humans less rational than lower animals? *Psychological bulletin*, *125*(5), 591.

Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The economic journal*, *99*(394), 116–131.

Baker, M. J. (2000). Writing a literature review. *The marketing review*, *1*(2), 219–247.

Bakos, Y. (1998). The emerging role of electronic marketplaces on the internet. *Communications of the ACM*, *41*(8), 35–42.

Barnes, S. B. (2006). A privacy paradox: Social networking in the united states. *First Monday*.

Baroni, L. A., Puska, A. A., de Castro Salgado, L. C., & Pereira, R. (2021). Dark patterns: Towards a socio-technical approach. *Proceedings of the XX Brazilian Symposium on Human Factors in Computing Systems*, 1–7.

Bechara, A. (2004). The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and cognition*, *55*(1), 30–40.

Berger, M., Greinacher, E., & Wolf, L. (2022). Digital nudging to promote energy conservation behavior–framing and default rules in a smart home app. *ECIS 2022 Research Papers*, *92*.

Bohnsack, R., & Liesner, M. M. (2019). What the hack? a growth hacking taxonomy and practical applications for firms. *Business horizons*, *62*(6), 799–818.

Bongard-Blanchy, K., Rossi, A., Rivas, S., Doublet, S., Koenig, V., & Lenzini, G. (2021). " i am definitely manipulated, even when i am aware of it. it's ridiculous!"-dark patterns from the end-user perspective. *Designing Interactive Systems Conference 2021*, 763–776.

Bösch, C., Erb, B., Kargl, F., Kopp, H., & Pfattheicher, S. (2016). Tales from the dark side: Privacy dark strategies and privacy dark patterns. *Proceedings on Privacy Enhancing Technologies*, *2016*(4), 237–254. https://doi.org/10.1515/popets-2016-0038

Boudreau, K. J. (2012). Let a thousand flowers bloom? an early look at large numbers of software app developers and patterns of innovation. *Organization Science*, *23*(5), 1409–1427.

Boudreau, K. J., & Hagiu, A. (2009). Platform rules: Multi-sided platforms as regulators. *Platforms, markets and innovation*, *1*, 163–191.

Brignull, H., Miquel, M., Rosenberg, J., & Offer, J. (2010). *Deceptive design user interfaces designed to trick people*. Retrieved July 7, 2022, from https://www.deceptive.design/

Brocke, J. v., Simons, A., Niehaves, B., Niehaves, B., Reimer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the giant: On the importance of rigour in documenting the literature search process. *ECIS 2009 Proceedings*, *161*.

Calo, R., & Rosenblat, A. (2017). The taking economy: Uber, information, and power. *Columbia Law Review*, *117*, 1623.

Chromik, M., Eiband, M., Völkel, S. T., & Buschek, D. (2019). Dark patterns of explainability, transparency, and user control for intelligent systems. *IUI workshops*, *2327*.

Constantiou, I. D., & Kallinikos, J. (2015). New games, new rules: Big data and the changing context of strategy. *Journal of Information Technology*, *30*(1), 44–57.

Conti, G., & Sobiesk, E. (2010). Malicious interface design: Exploiting the user. *Proceedings of the 19th international conference on World wide web*, 271–280.

Curchod, C., Patriotta, G., Cohen, L., & Neysen, N. (2020). Working for an algorithm: Power asymmetries and agency in online work settings. *Administrative Science Quarterly*, *65*(3), 644–676.

De Reuver, M., Sørensen, C., & Basole, R. C. (2018). The digital platform: A research agenda. *Journal of Information Technology*, *33*(2), 124–135.

Deng, X., Joshi, K. D., & Galliers, R. D. (2016). The duality of empowerment and marginalization in microtask crowdsourcing. *MIS quarterly*, *40*(2), 279–302.

Di Geronimo, L., Braz, L., Fregnan, E., Palomba, F., & Bacchelli, A. (2020). Ui dark patterns and where to find them: A study on mobile applications and user perception. *Proceedings of the 2020 CHI conference on human factors in computing systems*, 1–14.

Draper, N. A. (2017). From privacy pragmatist to privacy resigned: Challenging narratives of rational choice in digital privacy debates. *Policy & Internet*, *9*(2), 232–251.

Eaton, B., Elaluf-Calderwood, S., Sørensen, C., & Yoo, Y. (2015). Distributed tuning of boundary resources. *MIS quarterly*, *39*(1), 217–244.

Farrell, J., & Klemperer, P. (2007). Coordination and lock-in: Competition with switching costs and network effects. *Handbook of industrial organization*, *3*, 1967–2072.

Fischhoff, B. (2002). Heuristics and biases in application. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 730–748). Cambridge University Press. https://doi.org/10.1017/CBO9780511808098.043

Fogg, B. J. (2009). A behavior model for persuasive design. *Proceedings of the 4th international Conference on Persuasive Technology*, 1–7.

Galliers, R. D., Newell, S., Shanks, G., & Topi, H. (2017). Datification and its human, organizational and societal effects. *The Journal of Strategic Information Systems*, *26*(3), 185–190.

Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research policy*, *43*(7), 1239–1249.

Ghazawneh, A., & Henfridsson, O. (2015). A paradigmatic analysis of digital application marketplaces. *Journal of Information Technology*, *30*(3), 198–208.

Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008). Methods of data collection in qualitative research: Interviews and focus groups. *British dental journal*, *204*(6), 291–295.

Gray, C. M., Chen, J., Chivukula, S. S., & Qu, L. (2021). End user accounts of dark patterns as felt manipulation. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1–25.

Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The dark (patterns) side of ux design. *Proceedings of the 2018 CHI conference on human factors in computing systems*, 1–14.

Gulati, R., & Kletter, D. (2005). Shrinking core, expanding periphery: The relational architecture of high-performing organizations. *California Management Review*, *47*(3), 77–104.

Gulati, R., Puranam, P., & Tushman, M. (2012). Meta-organization design: Rethinking design in interorganizational and community contexts. *Strategic management journal*, *33*(6), 571–586.

Gunawan, J., Pradeep, A., Choffnes, D., Hartzog, W., & Wilson, C. (2021). A comparative study of dark patterns across web and mobile modalities. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1–29.

Henfridsson, O., Mathiassen, L., & Svahn, F. (2014). Managing technological change in the digital age: The role of architectural frames. *Journal of Information Technology*, *29*(1), 27–43.

Iansiti, M., & Levien, R. (2004). *The keystone advantage: What the new dynamics of business ecosystems mean for strategy, innovation, and sustainability*. Harvard Business Press.

Johnson, E. J., Bellman, S., & Lohse, G. L. (2002). Defaults, framing and privacy: Why opting in-opting out1. *Marketing letters*, *13*(1), 5–15.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kallinikos, J., Aaltonen, A., & Marton, A. (2013). The ambivalent ontology of digital artifacts. *Mis Quarterly*, 357–370.

Kant, I. (1870). *Grundlegung zur metaphysik der sitten* (Vol. 28). L. Heimann.

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, *14*(1), 366–410.

Kim, W. G., Pillai, S. G., Haldorai, K., Ahmad, W., et al. (2021). Dark patterns used by online travel agency websites. *Annals of tourism research*, *88*(100).

Kimppa, K. K. (2020). From just consequentialism to intentional consequentialism in computing. *ETHICOMP 2020*, 428.

Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & security*, *64*, 122–134.

Koontz, H. (1961). The management theory jungle. *Academy of Management journal*, *4*(3), 174–188.

Kuhn, K. M., & Maleki, A. (2017). Micro-entrepreneurs, dependent contractors, and instaserfs: Understanding online labor platform workforces. *Academy of Management Perspectives*, *31*(3), 183–200.

Lee, M. K., Kusbit, D., Metsky, E., & Dabbish, L. (2015). Working with machines: The impact of algorithmic and data-driven management on human workers. *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 1603–1612.

Luguri, J., & Strahilevitz, L. J. (2021). Shining a light on dark patterns. *Journal of Legal Analysis*, *13*(1), 43–109.

Maier, M., & Harr, R. (2020). Dark design patterns: An end-user perspective. *Human Technology*, *16*(2).

Manyika, J., Lund, S., Bughin, J., Robinson, K., Mischke, J., & Mahajan, D. (2018). *Independent work choice necessity and the gig economy* (tech. rep.). McKinsey Global Institute.

Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale: Findings from a crawl of 11k shopping websites. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–32.

Mathur, A., Kshirsagar, M., & Mayer, J. (2021). What makes a dark pattern... dark? design attributes, normative considerations, and measurement methods. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–18.

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

Mildner, T., & Savino, G.-L. (2021). Ethical user interfaces: Exploring the effects of dark patterns on facebook. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–7.

Moen, G. M., Ravna, A. K., & Myrstad, F. (2018). Deceived by design, how tech companies use dark patterns to discourage us from exercising our rights to privacy. *Norwegian Consumer Council Report*.

Möhlmann, M., Zalmanson, L., Henfridsson, O., & Gregory, R. W. (2021). Algorithmic management of work on online labor platforms: When matching meets control. *MIS Quarterly*, *45*(4).

Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (2015). Conducting semi-structured interviews. *Handbook of practical program evaluation*, *492*, 492.

Newell, S., & Marabelli, M. (2015). Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'. *The Journal of Strategic Information Systems*, *24*(1), 3–14.

Özdemir, . (2020). Digital nudges and dark patterns: The angels and the archfiends of digital communication. *Digital Scholarship in the Humanities*, *35*(2), 417–428.

Palmer, M. (2006). *Data is the new oil*. Retrieved June 4, 2022, from https://ana.blogs.com/maestros/2006/11/data%5C_is%5C_the%5C_new.html

Parker, G., Van Alstyne, M., & Choudary, S. P. (2016). *Platform revolution: How networked markets are transforming the economy and how to make them work for you*. WW Norton & Company.

Pavlou, P. A., & Dimoka, A. (2006). The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation. *Information systems research*, *17*(4), 392–414.

Raichle, M. E. (2010). Two views of brain function. *Trends in cognitive sciences*, *14*(4), 180–190.

Redström, J. (2006). Persuasive design: Fringes and foundations. *International Conference on Persuasive Technology*, 112–122.

Rolland, K. H., Mathiassen, L., & Rai, A. (2018). Managing digital platforms in user organizations: The interactions between digital options and digital debt. *Information Systems Research*, *29*(2), 419–443.

Rotella, P. (2012). *Is data the new oil?* Retrieved June 4, 2022, from https://www.forbes.com/sites/perryrotella/2012/04/02/is-data-the-new-oil/

Scheiber, N. (2017). How uber uses psychological tricks to push its drivers' buttons. *The New York Times*, *2*.

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*(1-2), 11–36.

Sherif, M. (1936). *The psychology of social norms*. Harper.

Sinnott-Armstrong, W. (2003). Consequentialism. *Stanford Encyclopedia of Philosophy*.

Stahl, B. C., Eden, G., Jirotka, M., & Coeckelbergh, M. (2014). From computer ethics to responsible research and innovation in ict: The transition of reference discourses informing ethics-related research in information systems. *Information & Management*, *51*(6), 810–818.

Täuscher, K., & Laudien, S. M. (2018). Understanding platform business models: A mixed methods study of marketplaces. *European Management Journal*, *36*(3), 319–329.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Tiwana, A. (2013). *Platform ecosystems: Aligning architecture, governance, and strategy*. Newnes.

Tiwana, A., & Konsynski, B. (2010). Complementarities between organizational it architecture and governance structure. *Information Systems Research*, *21*(2), 288–304.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, *185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (1985). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458.

Vallas, S., & Schor, J. B. (2020). What do platforms do? understanding the gig economy. *Annual Review of Sociology*, *46*(1), 273–294.

van Epps, E. M., Downs, J. S., & Loewenstein, G. (2016). Calorie label formats: Using numeric and traffic light calorie labels to reduce lunch calories. *Journal of Public Policy & Marketing*, *35*(1), 26–36.

Västfjäll, D., Slovic, P., Burns, W. J., Erlandsson, A., Koppel, L., Asutay, E., & Tinghög, G. (2016). The arithmetic of emotion: Integration of incidental and integral affect in judgments and decisions. *Frontiers in psychology*, *7*, 325.

Waldman, A. E. (2020). Cognitive biases, dark patterns, and the 'privacy paradox'. *Current opinion in psychology*, *31*, 105–109.

Wareham, J., Fox, P. B., & Cano Giner, J. L. (2014). Technology ecosystem governance. *Organization science*, *25*(4), 1195–1215.

Wason, P. C., & Evans, J. S. B. (1974). Dual processes in reasoning? *Cognition*, *3*(2), 141–154.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, xiii–xxiii.

Weinmann, M., Schneider, C., & Brocke, J. v. (2016). Digital nudging. *Business & Information Systems Engineering*, *58*(6), 433–436.

Woodcock, J., & Graham, M. (2019). *The gig economy: A critical introduction*. Polity.

Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). Research commentary–the new organizing logic of digital innovation: An agenda for information systems research. *Information systems research*, *21*(4), 724–735.

Zagal, J. P., Björk, S., & Lewis, C. (2013). Dark patterns in the design of games. *Foundations of Digital Games 2013*.

Appendices

Appendix A Dark Pattern Example of Uber

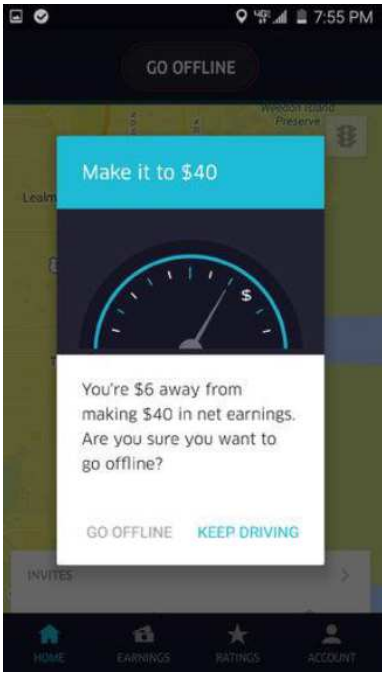

Figure 1: Taken from https://www.nytimes.com/interactive/2017/04/02/technology/uber-drivers-psychological-tricks.html (accessed at 04.06.2022)

Appendix B Flow Chart of Structured Literature Review

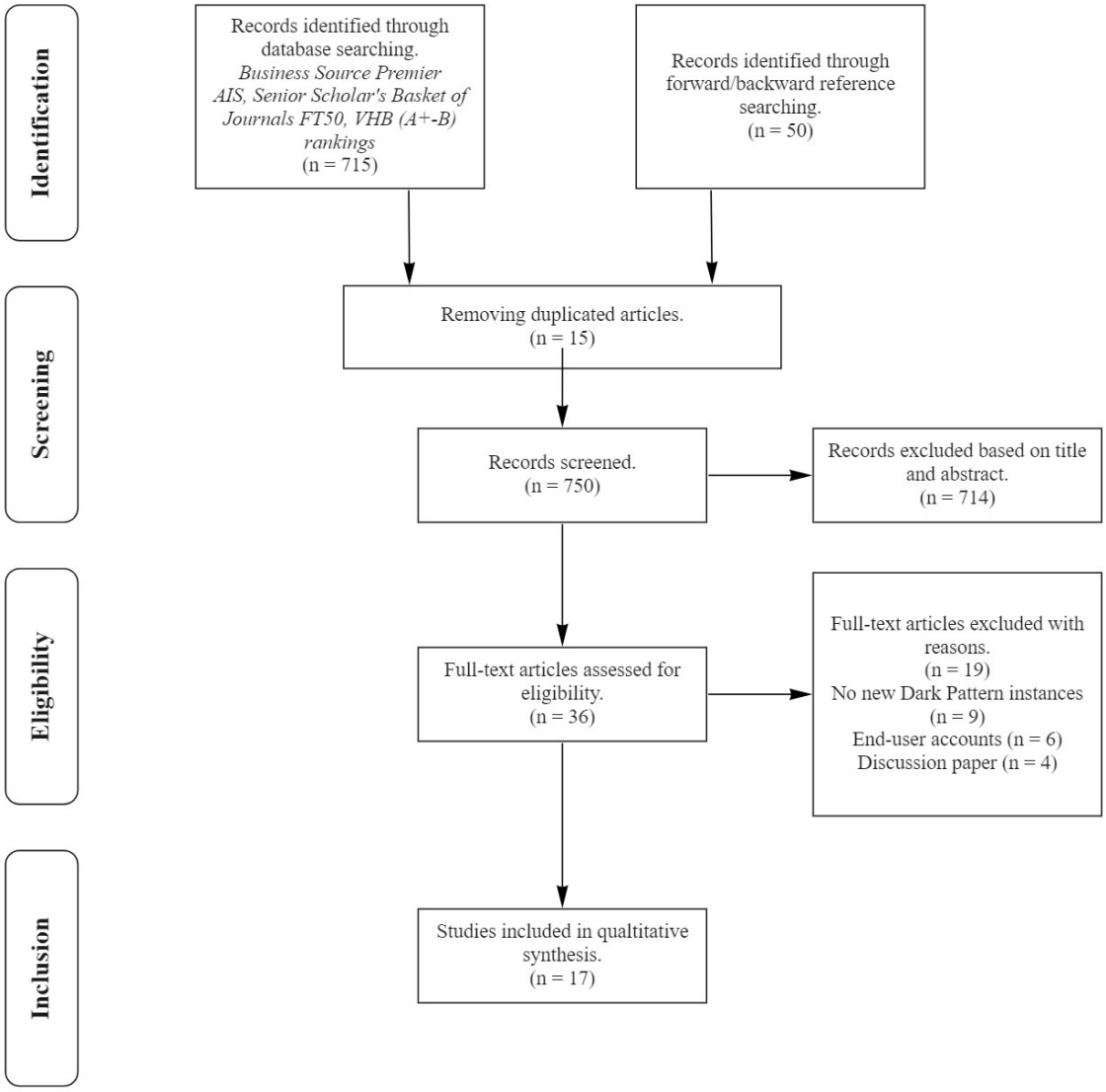

Figure 2: Flow Chart of Structured Literature Review

Appendix C Concept matrix

Articles	**Concepts**					
	Dark Pattern Strategies	Dark Pattern Types	Dark Pattern Characteristics	Cognitive Biases	Effectiveness	Usage Motivation
Baroni et al. (2021)		Roach motel	Socio-technical perspective			
Berger et al. (2022)				Default Rules & Framing		
Bösch et al. (2016)	Privacy context (maximize, publish, centralize, preserve, obscure, deny, violate, fake)	Bad Defaults, Privacy Zuckering, Forced Registration, Hidden Legalese Stipulations, Immortal Accounts, Address Book Leeching, Shadow User Profiles				

Continued on the next page

Table 2 (*continued*)

	Dark Pattern Strategies	Dark Pattern Types	Dark Pattern Characteristics	Cognitive Biases	Effectiveness	Usage Motivation
Chromik et al. (2019)	Nagging, Obstruction, Sneaking, Interface Interference, Forced Action				Lack of legislation	Legal profit increase

Continued on the next page

Table 2 (*continued*)

	Dark Pattern Strategies	Dark Pattern Types	Dark Pattern Characteristics	Cognitive Biases	Effectiveness	Usage Motivation
Conti and Sobiesk (2010)	Coercion, Confusion, Distraction, Exploiting Errors, Forced Work, Interruption, Manipulating Navigation, Obfuscation, Restricting Functionality, Shock, Trick	**Coercion** (Mandatory form field entries), **Confusion** (Incomprehensible questions), **Distraction** (Video/animation/blinking, color), **Exploiting Errors** (Typing errors), **Forced Work** (Difficult uninstall procedures), **Interruption** (Force viewing), **Manipulating Navigation** (Dead end trails), **Obfuscation** (Low contrast color scheme), **Restricting Functionality** (Omit necessary controls), **Shock** (Controversial content), **Trick** (Invisible installments, spoof content)				Revenue driven (increase brand recognition, gather personal information), others (shock, disgust, attack user)

Continued on the next page

Table 2 (*continued*)

	Dark Pattern Strategies	Dark Pattern Types	Dark Pattern Characteristics	Cognitive Biases	Effectiveness	Usage Motivation
Di Geronimo et al. (2020)	Nagging, Obstruction, Sneaking, Interface Interference, Forced Action	Extension of Aesthetic manipulation and Forced Action, **Obstruction** (Intermediate Currency, Price Comparison Prevention, Roach Motel), **Sneaking** (Bait and Switch, Hidden Costs, Sneak into Basket, Forced Continuity), **Interface Interference** (Hidden Information, Preselection, Aesthetic Manipulation, Toying with emotions, False Hierarchy, Disguised Ads, Trick Questions), **Forced Action** (Social Pyramid, Privacy Zuckering, Gamification)				

Continued on the next page

Table 2 (*continued*)

	Dark Pattern Strategies	Dark Pattern Types	Dark Pattern Characteristics	Cognitive Biases	Effectiveness	Usage Motivation
Gunawan et al. (2021)		Roach Motel, Intermediate Currency, Gamification, Forced Action, Privacy Zuckering, Hidden Costs	Technical perspective with modalities			

Continued on the next page

Table 2 (*continued*)

	Dark Pattern Strategies	Dark Pattern Types	Dark Pattern Characteristics	Cognitive Biases	Effectiveness	Usage Motivation
Gray et al. (2018)	Nagging, Obstruction, Sneaking, Interface Interference, Forced Action	**Obstruction** (Roach Motel, Price Comparison Prevention, Intermediate Currency), **Sneaking** (Forced Continuity, Hidden Costs, Sneak into Basket, Bait and Switch), **Interface Interference** (Hidden Information, Preselection, Aesthetic Manipulation, Toying with Emotion, False Hierarchy, Disguised Ad, Trick Question), **Forced Action** (Social Pyramid, Privacy Zuckering, Gamification)				

Continued on the next page

Table 2 (*continued*)

	Dark Pattern Strategies	Dark Pattern Types	Dark Pattern Characteristics	Cognitive Biases	Effectiveness	Usage Motivation
Kim et al. (2021)		False discount claims, hidden cost, activity notification		Anchoring Bias, Sunk Cost Fallacy, Bandwagon Effect		
Luguri and Strahilevitz (2021)	Nagging, Social Proof, Obstruction, Sneaking, Interface Interference, Forced Action, Scarcity, Urgency	Roach Motel, Sneak into Basket, Forced Continuity, Bait and Switch, Preselection, False Hierarchy, Trick Questions, Confirmshaming, High-demand Messages, Activity Message, Testimonial of Unknown Origin		Framing, Sunk Cost Fallacy, Anchoring	Forcing user to buy or disclose more products/information (undesired compulsion), cheap and effective sales tactics	Increase profits through additional purchases or data disclosure

Continued on the next page

Table 2 (*continued*)

	Dark Pattern Strategies	Dark Pattern Types	Dark Pattern Characteristics	Cognitive Biases	Effectiveness	Usage Motivation
Mathur et al. (2019)	Sneaking, Urgency, Misdirection, Social Proof, Scarcity, Obstruction, Forced Action	**Sneaking** (Sneak into Basket, Hidden Costs, Hidden Subscription), **Urgency** (Countdown Timer, Limited-time Message), **Misdirection** (Confirmshaming, Visual Interference, Trick Questions, Preselection), **Social Proof** (Activity Message, Testimonials of Unknown Origin), **Scarcity** (Low-stock Message, High-demand Message), **Obstruction** (Hard to Cancel), **Forced Action** (Forced Enrollment)	Asymmetric, covert, deceptive, information hiding, restrictive	Anchoring, Bandwagon Effect, Default Effect, Framing, Scarcity, Sunk Cost Fallacy		Trigger impulse buying

Continued on the next page

Table 2 (*continued*)

	Dark Pattern Strategies	Dark Pattern Types	Dark Pattern Characteristics	Cognitive Biases	Effectiveness	Usage Motivation
Mathur et al. (2021)			Disparate treatment			
Mildner and Savino (2021)	Interface Interference	Guided Setting				
Moen et al. (2018)		Default, Ease, Framing, Reward & Punishment, Forced Action, Timing		Hyperbolic Discounting, Anchoring		Customer data as main income source

Continued on the next page

Table 2 (*continued*)

	Dark Pattern Strategies	Dark Pattern Types	Dark Pattern Characteristics	Cognitive Biases	Effectiveness	Usage Motivation
Özdemir (2020)		Bait and Switch, Confirmshaming, Disguised Ads, Forced Continuity, Friend Spam, Hidden Costs, Misdirection, Price Comparison Prevention, Privacy Zuckering, Roach Motel, Sneak into Basket, Trick Questions		Social Norms, Anchoring		
Waldman (2020)			Anchoring, Framing, Hyperbolic Discounting, Default Effect			

Continued on the next page

Table 2 (*continued*)

	Dark Pattern Strategies	Dark Pattern Types	Dark Pattern Characteristics	Cognitive Biases	Effectiveness	Usage Motivation
Zagal et al. (2013)	Temporal, Monetary, Social-capital-based	Grinding, Playing by Appointment, Pay to Skip, Pre-delivered Content, Monetized Rivalries, Social Pyramid Schemes, Impersonation				

Table 2: Concept matrix

Appendix D Dark Pattern taxonomy

Strategy	Instance	Cognitive bias
Nagging		Framing Effect
Obstruction	Roach Motel	Sunk Cost Fallacy
	Price Comparison Prevention	Sunk Cost Fallacy
	Intermediate Currency	Hyperbolic Discounting & Sunk Cost Fallacy
	Immortal Accounts	Hyperbolic Discounting
Sneaking	Sneak into Basket	Default Effect
	Hidden Costs	Sunk Cost Fallacy
	Forced Continuity	Default Effect
	Bait and Switch	Framing Effect
Interface Interference	Aesthetic Manipulation	Anchoring & Framing Effect
	Preselection	Default Effect
	Hidden Information	Framing Effect
	Emotion Toying	Framing Effect
	False Hierarchy	Framing Effect
	Trick Questions	Default & Framing Effect
	Disguised Ads	Framing Effect
	Guided Setting	Framing effect
	Confirmshaming	Framing Effect
Forced Action	Social Pyramid	Framing Effect & Hyperbolic Discounting
	Privacy Zuckering	Framing Effect & Hyperbolic Discounting
	Gamification	Hyperbolic Discounting
	Pay to Skip	Hyperbolic Discounting
	Pre-delivered Content	Hyperbolic Discounting
	Pay to Win	Framing Effect & Hyperbolic Discounting
	Forced Registration	Hyperbolic Discounting
Scarcity	Low-stock Message	Scarcity Bias
	High-demand Message	Scarcity Bias

Continued on the next page

Table 3 (*continued*)

Strategy	**Instance**	**Cognitive bias**
Urgency	Countdown Timer	Scarcity Bias
	Limited-time Message	Scarcity Bias
Social Proof	Activity Messages	Bandwagon Effect
	Testimonials of Unknown Origins	Bandwagon Effect
	Impersonation	Bandwagon Effect

Table 3: Dark Pattern taxonomy

Appendix E Dark Pattern Interview Examples

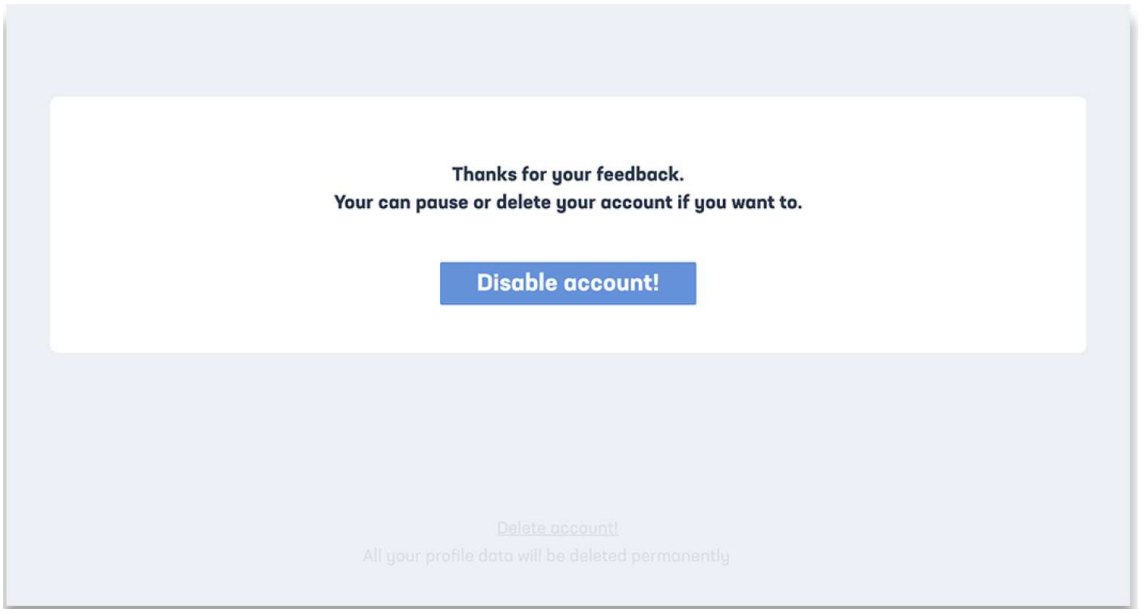

Figure 3: Hidden information & Trick Question Dark Pattern - Taken from: (Bongard-Blanchy et al., 2021)

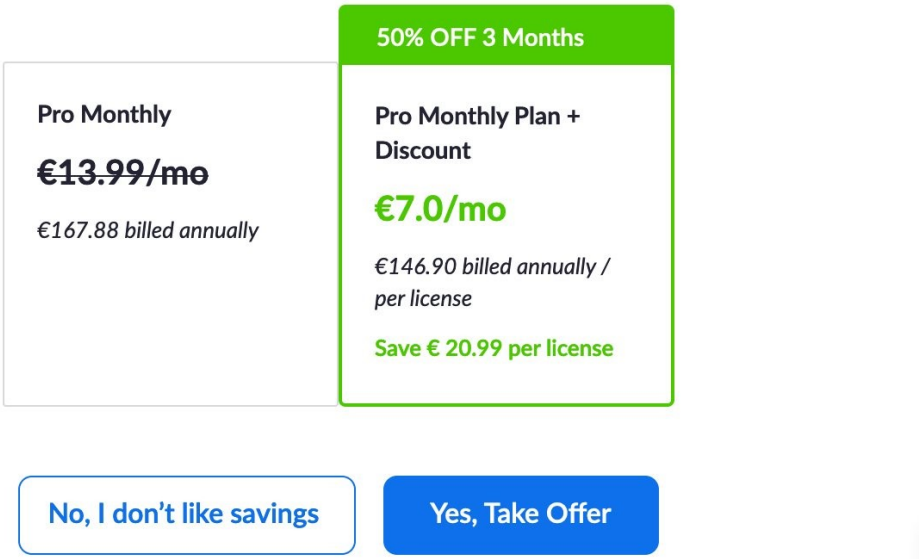

Figure 4: Aesthetic Manipulation & Confirmshaming Dark Pattern - Taken from: (Bongard-Blanchy et al., 2021)

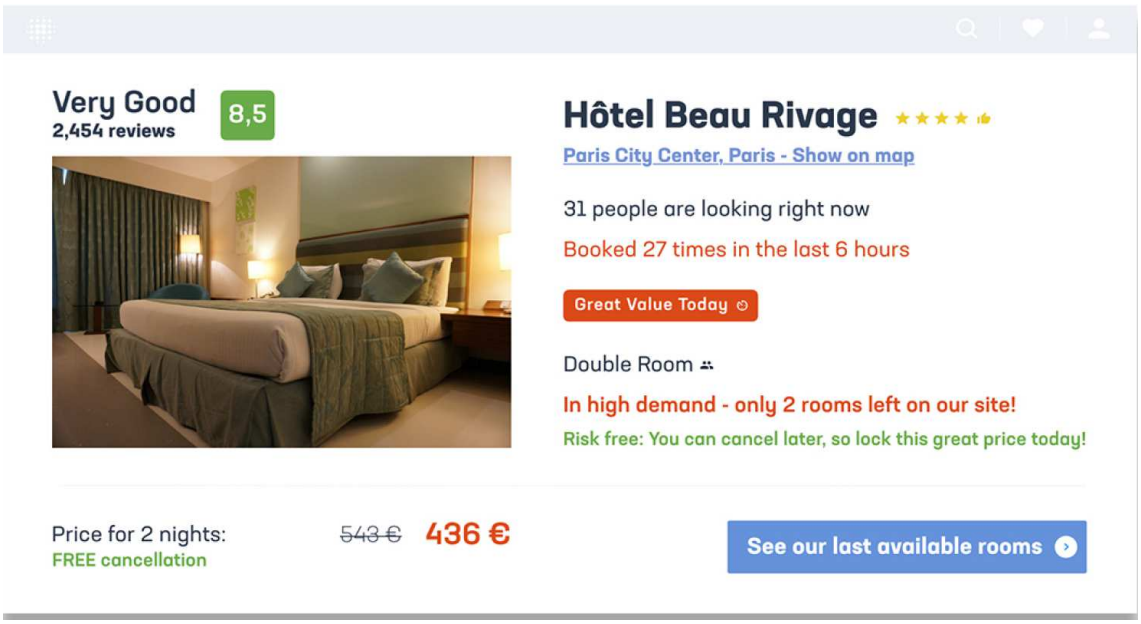

Figure 5: High-demand Message, Limited-time Message, Activity Message Dark Pattern - Taken from: (Bongard-Blanchy et al., 2021)

Appendix F Semi-structured Interview Guide

1. **Theme: Awareness, Harm and Prevalence of Dark Patterns**

   (a) What was the first thing that came into your mind about these Dark Patterns?

(b) Can you elaborate on the kind of emotions you are feeling when being exposed to these design interfaces? On which factors do these emotion depend?

(c) Could you immediately recognize these design interfaces as malicious intents or maybe you spotted it after some time passed?

(d) Would these Dark Patterns cause any harm for you and if what kind of harm?

(e) Where did you already find Dark Patterns?

2. **Theme: Worries through Manipulation of Dark Patterns**

(a) How do you usually react after you realize that you have been tricked into doing something you might not have done otherwise?

(b) Do you think it is your fault when something undesirable happens or do you think you are being taken advantage off?

3. **Theme: Acceptance of Dark Patterns**

(a) Do you excuse such manipulative behavior under certain circumstances?

(b) Are there examples when you desire such manipulative behavior?

(c) What do you think about using a free online service that collects and sells excessive amounts of personal data?

4. **Theme: Future Habits and Usage of Digital Services**

(a) Do you still use/would you still use the service in the future after experiencing Dark Patterns?

(b) When would you stop using a service that is deploying dark patterns?

(c) Is there any special kind of service that you would immediately leave after experiencing dark patterns?

5. **Theme: Countermeasures for Dark Patterns**

(a) What do you think, is it possible to fully avoid such manipulation?

(b) What might help you to avoid such manipulation?

6. **Theme: Company Perspective**

(a) What do you think about the motivation of companies employing such deceptive techniques?

(b) What is your opinion about what kind and size of companies are using Dark Patterns?

(c) Who is responsible in the company for deploying these design interfaces?