

Trivializing Values with Dark Patterns

An ethical analysis of the use of Dark Patterns by social media platforms and their effect on personal and societal values

By: Zachary S. Hankel

Date: July 16, 2024

Word count of main text (pp. 6-60): 23,922

Supervisor: dr. Y.J. Erden

Second Reader/Examiner: prof. dr. P.A.E. Brey

Final Master Thesis

MSc Philosophy of Science, Technology, and Society (PSTS)

University of Twente, Faculty of Behavioural, Management, and Social Sciences, Enschede, the Netherlands

Abstract

In this thesis I argue that social media platforms, through the use of manipulative software interfaces known as 'Dark Patterns', have the potential to interfere with an end user's ability to live by their values. Moreover, this thesis argues that, to the extent that Dark Patterns do interfere in this way, there is a further potential that these interfaces may interfere with society's ability to collaborate on the collective pursuit of goodness. Dark Patterns are understood to be digital interfaces which, by virtue of their design, influence an end user of some software into making online choices that are not necessarily in their best interest (Gray et al., 2018). I argue that even when a platform's values are not explicitly used to manipulate end users, the platform's values can be considered to be favored when the platform undermines the values of their users through manipulative tactics. A major aim of a given platform is to capture and keep a user's attention so they might turn that attention into behaviors which commercially benefit the platform. This thesis argues that platforms are quite successful in achieving this aim in part through the use of Dark Patterns. This process may simultaneously promote the platform's values while also undermining their users' values. This likely happens to the extent that Dark Patterns are systemically used which is argued to be significantly so. For those who wish to use their cognitive faculties of reflection and deliberation to help them live in accordance with their values, Dark Patterns may present a substantial obstacle. As a result, it is argued that there is good ethical reason to call for measures which curtail the use of Dark Patterns in software design.

Acknowledgements

My dearest thanks to the following supporters. I couldn't have done this without you.

dr. Y.J. Erden (Yj): Thank you for your patience, encouragement, and honesty in guiding me through this writing process. It's been a pleasure.

prof. dr. P.A.E. Brey (Philip): Thank you for your wisdom and willingness to help me brainstorm ideas.

dr. A. Henschke (Adam): Thank you for the conversations and for stepping in to help with my unique circumstances.

Jenni: Thank you for going on this journey with me, for your willingness to review my work, and for putting up with all of my ideas. <3

Jamie & Blake: Thank you for the games, lending your ears to my nonsense, and for being my connections to home.

Mom & Dad: Thank you for your love and encouragement and for taking the time to read my essays and letting me practice my presentation skills with you.

PSTS Class of 2022-24: Thank you for the fascinating conversations, the genuine interest, the insightful feedback, and the sharing of existential crises.

Table of Contents

Abstract	2
Acknowledgements	3
Table of Contents	4
Introduction.....	6
Chapter 1 - Values and Manipulation.....	9
Chapter 1 Introduction.....	9
Section 1.1. What are values?	9
Section 1.2. Values affect behavior	15
Section 1.3. Values take shape in practice	19
Section 1.4. Manipulation may influence values.....	22
Chapter 2 - Values and Dark Patterns.....	26
Chapter 2 Introduction.....	26
Section 2.1. A brief history of Dark Patterns	27
Section 2.2. How Dark Patterns manipulate.....	33
Section 2.3. The systemic nature of Dark Patterns	36
Section 2.4. Dark Patterns may influence a person’s values.....	39
Chapter 3 - Dark Patterns and Ethics.....	43
Chapter 3 Introduction.....	43
Section 3.1. Dark Patterns can disrupt ethical decision making	44
Section 3.2. Dark Patterns can promote unethical values	47
Section 3.3. Dark Patterns can hinder ethical growth.....	49
Section 3.4. Some ethical suggestions	52
3.4.1. Resistance	52
3.4.2. Awareness and Transparency.....	54
3.4.3. Alternative design patterns	55
Conclusion	57
References	61
Appendices.....	70
Appendix A - Screenshots of Dark Patterns.....	70
Examples from Gray et al. (2018)	70
Examples from Luguri & Strahilevitz (2021)	72
Examples from Mildner et al. (2023).....	73

An example from Nouwens et al. (2020) 74

Appendix B - Alternative Tables..... 77

Introduction

In this thesis I argue that social media platforms, through the use of software interfaces known as 'Dark Patterns', have the potential to interfere with an end user's ability to live by their values. Borrowing from discussions of values in the pragmatist tradition, I define values as conceptions of worth regarding goodness. Moreover, this thesis argues that, to the extent that Dark Patterns do interfere in this way, there is a further potential that these interfaces also interfere with society's ability to collaborate on the collective pursuit of goodness. Dark Patterns are understood to be digital interfaces which, by virtue of their design, manipulatively influence an end user of some software into making online choices that are not necessarily in their best interest (Gray et al., 2018). These interfaces are considered manipulative due to their tendency, through the presentation of on-screen options, to undermine end users' considerations and promote options which are more beneficial to the platform utilizing said interface. Online social media platforms, which this thesis sometimes abbreviates to 'platforms.', are examined as online virtual spaces engaged through a website or app where end users can socially connect to create, share, consume, and engage with content. With the increasing prevalence of online social media platforms and their industry of digital influence, I propose that platforms are a space that currently deserve critical attention (Giraldo-Luque et al., 2020; Borchers & Enke, 2022). With these ideas in mind, this thesis asks the following research question: 1. Why, how, and to what extent are Dark Patterns utilized by an online social media platform to manipulate end-user behavior and values in the service of that platform's commercial interest and what may be the harm of such manipulation?

To answer this question, the relationship between manipulation and values must first be examined. To that end, this thesis will first seek to answer the following research sub-question: 1.1. How, if at all, can manipulation performed by an online social media platform change an end-user's values? As indications of what matters to people or ideals that someone might aspire towards, values may be connected to what people think is right and good (Singer, 2024). As a result, if someone wants to make decisions that matter to them, then their values might be a good place to begin thinking about those decisions. If someone can make a decision that aligns with their values, they might reasonably expect the results of that decision to advance the kind of goodness their values express. However, this thesis argues that personal values are not necessarily static and may be subject to change (Dewey and Tufts, 1936). To the extent that they are subject to change, values might be susceptible to undue influence.

To continue answering the primary research question, the potential for a digital artifact to be used to manipulate must also be examined. To that end, this thesis will seek to answer the following research sub-question: 1.2. How might manipulation be designed in a technical artifact, and what might that mean for any values implicated through the use of that artifact? This thesis argues that Dark Patterns are digital artifacts utilized by platforms which manipulate end user's into making decisions that may align more with what the platform values and less with what the end user values (Susser et al., 2019; Luguri and Strahilevitz, 2021). This thesis details how Dark Patterns came about, how they are used, and examines their effectiveness before tying back to the previous sub-question. Once Dark Patterns have been discussed in detail, it will be possible to demonstrate why their manipulateness may be especially concerning for someone who wishes to live by their values.

Finally, to finish answering the primary research question, this thesis will examine the ethical aspects of the effect of Dark Patterns on values. To that end, this thesis will seek to answer the following research sub-question: 1.3 What are the ethical considerations of any potential harmful consequences of a manipulative digital artifact like Dark Patterns? This thesis argues that when an individual acts in accordance with any values, particularly in a public setting, then those values are likely to have some impact on others in society (Wieczorek, 2023). The people involved in these interactions may have different ideas of what matters to them or what they value. In this thesis I argue that, in order to make collaborative moral or ethical decisions, the people involved likely need to be able to understand and communicate the values they have as well as to be receptive and understanding towards the values of others. If people are less able to reflect and deliberate on what they themselves think is right or good, then they may also be less able to collaborate on right or good decisions. Collaborating on right or good decisions is conducive for making ethical decisions, particularly those that have a societal impact. To the extent that Dark Patterns interfere with an end user's ability to live by their values, they might also be interfering with the ability for society to progress ethically.

The methodology of this thesis is predominantly a conceptual analysis combined with secondary research. The main concept that will be analyzed is that of values which will be argued to be ways of conceptualizing 'the good' that often take shape in practice. The concept of manipulation will also be analyzed as the act of unduly influencing someone's decision in such a way that

undermines that person's judgment. These concepts will be developed through secondary research performed through an interdisciplinary approach.

The approach of this thesis is an interdisciplinary one that draws from pragmatism, ethics, psychology, user experience (UX) design, and human computer interaction (HCI). The argument developed will utilize the data and conclusions from empirical studies performed in HCI and UX which seek to understand Dark Patterns and their effects on end users. Since these studies make a connection between HCI/UX and certain theories in cognitive psychology, this thesis will also draw on studies in psychology. Due to the conceptual analysis of values and manipulation, a link is proposed between Dark Patterns and values which is, namely, that Dark Patterns may influence end user practice(s) such that they may influence end user values. Furthermore, the empirical studies in HCI/UX that this thesis utilizes make ethical claims which will be analyzed using the link this thesis proposes.

Chapter 1 seeks to answer research sub-question 1.1 through a conceptual analysis of values using a predominantly pragmatic approach, otherwise known as practice theory. Pragmatism asserts that in order to judge and reflect on right action, one must examine what happens in practice. Furthermore, pragmatism often understands 'meaning' as "the totality of conceivable [(i.e. able to be conceived)] practical consequences" and so there is an emphasis on the consequences of practice (Hickman et al., 2011, p. 186-187). Chapter 2 seeks to answer sub-question 1.2 by performing secondary research on Dark Patterns and continues with the conceptual analysis of their interaction with values. Dark Patterns are manipulative forms of design patterns which are repeatable software solutions in UX/HCI aimed at making software easy to understand and use (Zimmer, 1995). A hallmark of Dark Patterns is their exploitation of cognitive biases and, because of the conceptual analysis performed in chapter 1, a connection is made in cognitive psychology between values and Dark Patterns. Chapter 3 seeks to answer sub-question 1.3 by exploring this connection in detail, analyzing its ethical consequences, and proposing some solutions.

Chapter 1 - Values and Manipulation

Chapter 1 Introduction

This chapter will make arguments for how manipulation can change a person's values. To do so, I will discuss what values are and how they can function in someone's life, then I will discuss what manipulation is and how it can interact with values to potentially change them. In terms of what values are, the definition I put forward in this chapter is largely informed by the pragmatic approach of John Dewey (1922b) and other authors who have reflected on his work. That definition, namely, is that values are *conceptions of worth regarding goodness*. The first section of this chapter will break this definition down into its key terms and defend why this definition is a reasonable one. This discussion will emphasize the role of human fallibility among other complexities in the conceiving of values (Kaag, 2013). The second section will explain how values are used in everyday practice. Depending on one's attitude towards them, values tend to be persistent concepts that can guide the decisions of both individuals and groups (Hickman et al., 2011). Yet, simultaneously, values can be rather flexible concepts in the ways they are applied to various contexts (Wieczorek, 2023). The third section will then cover how values are formed. Consistent with the pragmatic approach of this thesis, values are not only used in practice but may take shape through practice. How values are thought of and used in practice is influenced by the members of a given society or culture (Le Dantec et al., 2009; Boenink and Kudina, 2020). Section four will end the chapter by expanding on how values are vulnerable to manipulation. This final section will first explain what manipulation is before detailing how it might be used to influence a person's values (Susser et al., 2019).

Section 1.1. What are values?

This section will argue for the definition of values as *conceptions of worth regarding goodness*. This definition is comprised of depictions of values in the pragmatic tradition along with discussions in the Oxford Handbook of Value Theory (Dewey, 1922b; Hirose & Olson, 2015). The key words to be examined here are conceptions, worth, and goodness. I will elaborate on these terms starting with goodness and working backwards through my definition. Goodness will be examined instrumentally (good for something else) and intrinsically (good for its own sake) (Korsgaard, 1983). Predicated on instrumental and intrinsic conceptions of goodness, worth will be examined as a general attitude that may lead to more specific conative attitudes, or attitudes

of what to do regarding an object of value (Svavarsdóttir, 2014). Finally, values will be discussed as conceptions that may be implicit/unconsidered or explicitly considered (Dewey, 1925). Judgment need not be performed for values to influence a person's actions, but judgment is likely necessary if one wishes to understand a given value and intentionally make decisions with it. In this way, how values appear in conceiving or thinking follows a dual process theory of human cognition. Dual process theory holds that cognitive processing often flows in two different modes: one automatic, the other deliberate (Frankish, 2010). Moreover, conceptions are necessarily fallible ideas and are subject to revision (Grenberg, 2007; Kaag, 2013).

Values in philosophy are usually “associated with what is good” (Van De Poel, 2020, p. 302). What a specific value means typically refers to what makes something good or what something is good for. A beautiful sunset, for instance, suggests that this sunset is good for someone who wishes to experience beauty. An efficient tool, on the other hand, suggests that the tool is good for accomplishing some tasks efficiently. These examples depict two different notions of the term ‘good for.’ Where the former depicts beauty as good for its own sake *qua* the experience of beauty, the latter depicts efficiency as good for something else. In axiology (or value theory), the former is typically called ‘final’ or ‘intrinsic’ value while the latter is called ‘instrumental’ value. This distinction is often made when discussing goodness where ‘good’ can mean some ideal end or it can refer to some useful means to an end (Korsgaard, 1983).

The ways values relate to goodness often follows this pattern of instrumental vs. final in value theory (Schroeder, 2021). Instrumental values are those which are good for the sake of something else while final values are those which have been judged to be good for their own sake (Korsgaard, 1983). In both cases, values relate to goodness as the answer to the query “what makes it good?”. The ‘what’ of a value is itself in the case of intrinsic value and it is how it contributes to some other end in the case of an instrumental value (Dewey, 1922b). To return to the examples of beauty and efficiency, beauty might be considered what makes some end good. Beauty might be thought of as something worth pursuing because there is something intrinsically good about beauty that makes it worth experiencing. Efficiency describes a good way a given entity might contribute to some end. There may not be anything intrinsically good about efficiency that makes it worth experiencing, but rather it is good if one wants to, for example, perform some tasks effectively. The kind or degree of goodness a given value might be ‘good for’ can be depicted in terms of what it is worth.

Dewey (1922b) considers 'worth' to be "the tribute paid by reason to value" (p. 351). In other words, to explain what something is worth is to attempt to reason why one should value it. If I were to ask, 'what is that worth to you?', it is reasonable to think that I expect you to quantify and/or qualify your desire for whatever 'that' is. In the Oxford Handbook of Value Theory, Pattanaik & Xu (2015) demonstrate this concept with the following thought experiment. Imagine you are presented with two situations and asked in which situation you would have a greater value of freedom. In situation A you are asked what you would like for dessert and presented with two choices: a fresh apple or nothing at all. In situation B you are asked the same thing and given the same choices with one addition: a rotten apple. Pattanaik & Xu (2015) argue that, to the extent that freedom in some way depends on the number of available options, a person might reasonably evaluate situation B to have greater freedom. However, it does not necessarily follow that the value of this freedom is necessarily greater as well. One might reasonably say "but this extra freedom is not worth anything to me" since it is highly unlikely anyone would choose a rotten apple especially if a fresh one is available (Pattanaik & Xu, 2015, p. 367). The sort of freedom to eat a rotten apple for dessert when a perfectly good apple is available is likely not worth very much, i.e., not very valuable. What this thought experiment demonstrates is that, whether freedom is valued instrumentally or intrinsically, it may be possible to quantify or qualify its value. The same can be considered for values generally.

The quantification/qualification of value may also be understood in terms of a general attitude towards the person, place, thing, or idea desired. A general attitude towards a place of historical significance, for instance, might be qualified as an allure of that place as a location important in history. This general attitude may help explain more specific emotional or motivational attitudes concerning what to do about the person, place, thing, or idea (Svavarsdóttir, 2014). For example, if an old building somewhere were discovered to have some great historical value, it may engender the notion that the building should be preserved. Furthermore, it may inspire an attitude of teaching history, prompting the construction of an educational repository such as a museum. This kind of attitude is also called a 'pro-attitude' or a fitting attitude in value theory which is the concept that values may carry certain attitudes that fit one's understanding of the value at hand (Rabinowicz & Rønnow-Rasmussen, 2004). There are some unresolved questions concerning this idea. For example, are pro-attitudes inspired by values or assigned to them? These complexities can make it difficult to discern the exact relationship between values and specific attitudes concerning those values. Still, it may be the case that pro-attitudes are at least a good indication of a respective value.

An emphasis in the pragmatist tradition is that value “is to be found in present experiences” as well as directed towards the future in the sense that values can be an assertion or prediction of what something may be (Hickman et al., 2011, p. 188; Gouinlock, 1978). To claim that something is ‘worth it’ is akin to claiming that whatever ‘it’ is ought to be desired. Whether something is desired because it is valuable or whether something is valuable because it desired is a contested debate in value theory, but no matter which side someone falls on in that debate it seems likely that there is a strong connection between desire and values (Oddie, 2015). If something is desirable, it may be thought to be so presently and predicted to remain so in the future. As conceptions of worth, values can be considered to be present and future oriented in that they indicate what one may do with the subject or object of value (Gouinlock, 1978). If a tool is fittingly described as efficient, then it must be so for a present or future task that it is made to perform. To lose its efficiency would be to lose this particular value and it may no longer be used to complete any tasks in an efficient manner unless it regained that value.

Worth and goodness are related but distinct concepts that are not necessarily equivalent. It is possible to perceive something as having worth without thinking of it as being good and vice versa. For example, one may recognize that others believe something is good while they themselves have no such belief. Still, knowing that others find this something to be good allows one to bestow upon it a sense of worth, generally speaking. However, when nothing good can be thought of regarding a certain something for anyone in any context, it might reasonably be evaluated to be worthless. Something may also be thought of as good but be difficult to quantify with a sense of worth, as may be the case with a sunset. However, even here it is reasonable to recognize that the goodness of a sunset, for example, may be something that someone finds worthwhile. Even though the worth of something may be difficult to pin down, one can still recognize that it can have worth.

Depicting value in terms of worth regarding goodness accommodates an understanding of how that value may matter or what meaning it might entail. The particular way a given value is good suggests how that value might be meaningful. If one desires to have the particular goodness of a given value in their life, then the way it is good or what it is worth suggests how that value may interact in their life. If beauty, for instance, is good to experience, then it is reasonable to think that it is worth the effort to experience. Worthiness suggests a fitting attitude, in this case, one of spending effort. Worthiness and meaningfulness are related concepts (Betzler, 2019). If I desire

to experience a beauty that I believe is worth experiencing, then that means I will likely need to take the time or make the effort to experience it if I am able. I might understand the meaning of beauty to be something that, in part, is good to experience to the point of being worthy of my time and effort in certain contexts. Of course, beauty means much more than merely being good enough to experience, but the desire to experience beauty may help to explain the motivation to get out of my chair to see a sunset.

Values may be best understood as conceptions, i.e., mental representations or abstractions. Values are not tangible objects but abstract principles which exist in one's mind that point to and qualify aspects of goodness. As conceptions, the worth of a value in one's understanding is necessarily an estimate or hypothesis (Dewey, 1922b, Kaag, 2013). The reason for this, I argue, is that what a given value means to someone is likely not something they can express fully. The pragmatist depicts meaning in terms of the totality or summation of conceivable consequences (Hickman et al., 2011). In other words, what something means is what happens or could happen to, within, and because of that thing over time. This is a tall order for complex concepts like values and, as a result, a value's full conceivable consequences or meaning is likely very difficult to obtain fully. This view is supported by the fact that many debates about goodness remain unsettled in axiology, as this section has touched on (Schroeder, 2021). Therefore, any current depiction of worth likely cannot fully capture every sense of what it might be worth. Moreover, what a value might be worth may not be something actively depicted at all which may be the case for any value held implicitly. As a result, the term 'conceptions' attempts to capture these complexities suggesting that values are general ideas which are imperfect or unfinished.

Values as conceptions can be understood as either implicit or explicit ideas. An implicit value may be an unconsidered preference, but one which still generally captures a sense of goodness and worth as discussed previously. What something is worth can be implicitly thought and felt. For an implicit value, one might be automatically conceiving it as feeling or seeming good without being able to express why they perceive it that way. For example, if some artisan personally holds the value of efficiency implicitly, then they may gravitate towards using efficient tools without knowing precisely why. One might treat something as valuable without judgment (Dewey, 1922b). Once one makes the attempt to understand a value they may hold, that value becomes an explicit conception. Explicit conceptions, or judgments, are formed through reflection and deliberation. Once made explicit, an individual who holds a given value will likely be able to say something

explicit about it, i.e., they may be able to formulate a judgment about what that value means to them.

Thinking of values as conceptions which might be either implicit or explicit largely follows the same line of reasoning as dual process theory. Dual process theory comprises a family of theories of cognition which allege that human thinking processes predominantly fall into two different modes (Frankish, 2010). The first mode, known as Type 1 or System 1, describes cognitive processes which are unconscious, implicit, automatic, fast, and readily available. Types of System 1 cognitive processes include biases, gut reactions, and impulses. System 1 thinking is advantageous in that it allows for quick decision making and seems relatively undemanding cognitively. Implicit values as automatic conceptions may follow a similar course to System 1 cognition and so may be more unconscious, fast, and readily available. The second mode, known as Type 2 or System 2, describes cognitive processes which are conscious, explicit, controlled, slow, and analytic. Types of System 2 cognitive processes include logical thinking, reflection, and deliberation. System 2 thinking is advantageous in that it allows for considered decision making though it is fairly cognitively taxing as it requires concentration and discipline. Explicit values as considered conceptions may follow a similar course to System 2 cognition and so may be more conscious, slow, and controlled.

Finally, before moving on to the next section, it is important to emphasize the role of reflection in the process of valuation, or the conceptualization of values. In cognition, both implicit and explicit processes are prone to failure (Frankish, 2010). Likewise, one's pursuit of goodness is likely "an ever-continuing and fallible process, never fully achieved, subject to continual review and refinement, and, quite possibly, to failure" (Grenberg, 2007, p. 94). The propensity to failure may be a source of harm and wrongdoing, but it can also be an opportunity to reflect and improve (Dewey, 1922b; Kaag, 2013). Reflection can be considered a conscious attempt to assess a situation, seek to understand why that situation unfolded as it did, and improve one's judgment accordingly (Dewey and Tufts, 1936). For example, if I were to try to climb a mountain that turned out to be too difficult for me to complete, if I wish to reattempt to get to the top (perhaps because I value the achievement) then it may help me to reflect on what went wrong so that I may improve my next attempt. As such, reflection can be considered to be a System 2 sort of thinking as it is a conscious and deliberate cognitive process. Reflection of values also emphasizes a significant role that values can play in a person's life which is, namely, to affect one's behavior which is where the next section shall turn to.

Section 1.2. Values affect behavior

This section will argue that the attitudes that values engender indicate an idea of what to do with the object or subject of value. To support this idea, I will begin by exploring values as consistent yet flexible concepts. They are consistent, not in a fixed or permanent sense, but in the ways that they indicate consistent attitudes with regard to overall goals. This allows values to “endure from situation to situation” as their relevance may be understood to be widely applicable (Hickman et al., 2011, p. 179). However, to be applicable from situation to situation, values likely must be flexible enough to account for the relevant differences between situations if they are to adequately guide one’s behavior (Wieczorek, 2023). If this assessment is correct, then the appropriate view of any given value is one that is not so narrow or so fixed as to be impossible to fit into any specific application nor too broad or indefinite so as to become meaningless or worthless.

Once the scope of a given value is roughly conceived, if one wishes to act in a manner consistent with the values they hold, then one’s attitudes can allow them to proceed with a consistent attentiveness to the subject or object of value. I argue that this attentiveness provides one with focus as they make value-congruent decisions, i.e., decisions in line with the values they hold. This perspective is one that has recently found support in empirical research which is discussed in detail later on in this section (Sagiv and Roccas, 2021; Frömer & Shenhav, 2022). With one’s goal in focus, guided by the attitudes of any relevant values, an individual might be able to bring a value-informed decision to fruition. Furthermore, it is possible that an agent may find a way to get someone else to treat some object as valuable in a particular way without the latter making any judgment themselves. As a result, if it is possible to integrate a given value and the attitudes it engenders within an artifact or process then that value might still guide someone’s behavior (Klenk, 2021).

As argued in the previous section, values indicate appropriate or fitting conative attitudes. An historical building indicates potential ways to treat it, an efficient tool suggests how to use it, and a beautiful sunset suggests a kind of experience to expect from it. These attitudes are appropriate in that they make sense given the content of the value at hand (Svavarsdóttir, 2014). Some attitudes may make more or less sense given certain circumstances and are thus more or less appropriate. As discussed in the previous section, something that has value may reasonably be ascribed a sense of worth both quantitatively and/or qualitatively. In other words, the object of a value reasonably indicates one or more pro-attitudes, i.e., attitudes that advance said value.

These attitudes are not normatively required for someone to hold a relevant value, nor does the worthiness of a given value necessarily depend on someone's attitude, but rather and more simply, values can be suggestive of certain pro-attitudes (Svavarsdóttir, 2014). For example, a pro-attitude of minimizing waste may be reasonably appropriate regarding the value of sustainability, but it is perhaps not as reasonable regarding the value of pleasure. Though minimizing waste might be pleasurable in some cases, it is not necessarily so and in some cases it might be distinctly unpleasant. However, minimizing waste is more likely to be applicable in circumstances where sustainability is a relevant concern or value. Thus, the pro-attitudes of values can be relatively applicable in different contexts.

Before detailing how a value might be stable and flexible, I wish to assert that these concepts are not contradictory. Many systems, disciplines, and institutions must be both relatively stable and flexible to adequately function. Architecture, for example, requires enough stability in design so that a building in question does not topple, but it must also be flexible enough so that the building can absorb the force of the elements. Education must be stable enough so that its material is coherent, but flexible enough so that it might adapt to advancements in science, technology, and society. A democratic government must be stable enough so that law and order can be upheld but flexible enough to allow each citizen the freedom to be themselves. I propose axiology is no different. Stability and flexibility may be thought of as two ends of a spectrum and each value might be thought to fall somewhere on that spectrum relative to the attitudes it engenders. As a result, it is sensible to expect a certain consistency with values. Such consistency must be firm enough so that a value can be recognizable as one and the same between contexts but malleable enough for appropriate adjustments as the situation demands.

So then, I first wish to consider values as relatively stable concepts. I do not mean to support an objectivist account of values; I am not supposing that any value necessarily has a core meaning which transcends the situations in which it might be used. There are limits to this stability and given enough time, the meanings of a given value as we understand it today may evolve into something completely different. Instead, what I am proposing is that values may be understood to have core meanings as they are used in discourse between agents who understand one another. This form of stability allows for the evolution of a given value over time, but accounts for how someone might reasonably expect a value that was fitting in the past to be relevant for a similar situation (Boenink and Kudina, 2020). There is evidence in psychology that values operate like this on a cognitive level. Sagiv and Roccas (2021) found that adults tended to report a similar

level of importance across situations for a value they held. Self-reported behavior and peer-reported behavior suggested that an individual's values may be upheld over the course of months and even years. Enough stability is necessary to the extent that anyone can make a decision that is in line with their values (Hickman et al., 2011). If values were so unstable as to mean something completely different from moment to moment, there may be no way to predict what a value might mean next, let alone how to act in accordance with that value.

Values also need a level of flexibility if they are to endure with changing circumstances; a new circumstance should not necessarily or automatically render a relevant value obsolete. There may be good reasons to continue desiring whatever benefits a value might afford despite new evidence that puts the old understanding in question as is. An example of this can be found in an ongoing conversation on techno-moral disruption. Nickel et al. (2022) discuss how mechanical ventilators changed the medical understanding of death in the 1950's as well as the value of what can be called the respect for a patient's body. Human organs can only legally be removed from a corpse for the purpose of organ transplant in the EU and U.S. However, mechanical ventilators were able to keep a body alive indefinitely which challenged the conventional understanding of death and corpses. Suddenly, a patient could be something between life and death. Despite this advancement in medicine blurring the medical distinction between life and death, respect for a patient's body may reasonably remain important. There can still be good reasons to hold on to values despite changing conditions (Nickel et al., 2022). A change in circumstances can be such that a certain understanding of a value is no longer wholly fitting. If that value is to remain valuable then it must be flexible enough to adapt to the change in circumstance.

In summary, a value must be stable enough so that one might consistently conceive of it but it must also be flexible enough so their conception can adapt to new relevant situations. An explicit depiction of a value's stability and flexibility can also be thought of as the appropriateness of one's conceptions of worth, or attitudes, brought into focus. To continue with the previous example, respect for a patient's body must be consistent enough that one can recognize when that respect is violated. It must also be flexible enough that one can recognize the many different situations such respect might be warranted or not. These kinds of recognition requires focus or attentiveness. Though I am proposing that stability and flexibility are qualities of all values, to be able to focus on a given value likely requires explicit attention.

The attention someone pays to their values can help them make decisions guided by said values. One's attention towards a given value can be understood to be in line with their relevant attitudes (Svavarsdóttir, 2014). For example, if someone has an attitude of preservation regarding historical objects, then the attention they pay towards an historical object will likely concern its preservation. One's attitudes in focus may therefore be understood to be related to their attentiveness. If one is able to focus on their concern for a specific historical object, they may be able to deliberate on methods of how to preserve said object. In other words, values might be harnessed to intentionally affect one's behavior. This mechanism of how attention might help to transform a conception of value into action also has some support in empirical research such as psychology and neuroscience. Sagiv & Roccas (2021) propose that, insofar as one's attitudes concerning a value are goal oriented, they may be more willing to pay attention to that value in service of that goal. Frömer & Shenhav (2022), propose that attention spent on a value may help one to sample that value and get acquainted with it, which may also help in the consideration of value-congruent choices. In other words, these studies suggest that paying attention to values may aid one in making value-informed decisions.

This section has sought to trace a conceivable path from values to decisions and action. It may not be the only path, but it is one that explains how a perspective of what is good may serve as a guide for decision making. Furthermore, it helps to make sense of another phenomenon: how someone might guide the decisions of others using values. If values do inspire certain attitudes, and those attitudes suggest a specific kind of attention, then if a given value grabs someone's attention, either implicitly or explicitly, it might be of that specific kind. Recalling the example of the efficient tool, the way that it is efficient may reasonably grab someone's attention such that they find an attitude of efficient use, even if they don't fully comprehend it. The specific kind of attention likely does not appear out of thin air. It might be subtly informed by that value's deliberate use in social interactions or things which advance that value's conception of worth (Klenk, 2021). People who may not have an explicit conception of that value may still be exposed to it through such means. As a result, someone may attend to a value implicitly or automatically, as a preference or bias for example. Whether implicitly or explicitly, one's attitude concerning a given value can be informed by how it is used which may reasonably lead to value-congruent actions. This concept will be detailed in detail in the next section.

Section 1.3. Values take shape in practice

As argued, one's attitude(s) towards a given value gives an indication of what someone may wish to do concerning a subject or object which carries such value. This section will argue that acting in accordance with a given value, whether in an actual action or an imagined one, is an enactment or performance of said value. If an enactment of something allows one to experience that thing to some extent, then the same must be true of values. This experience of enactment may be what enables one to reflect on values and make changes to their understanding as they see fit (Dewey and Tufts, 1936). Thus, this section argues that a value's enactment is not a 'final say' as what happens after a value's moment of enactment continues to shape its meaning in a given culture or society. As such, this section will argue that values are continually constructed in practice as their ongoing evaluation is used to resolve problematic situations, or conflicts (Gouinlock, 1978; Hickman et al., 2011). In accordance with an everyday use which continually shapes them, this section will propose that values are predominantly formed in organic, bottom-up processes where 'ordinary' people conceive of what values are worth and how they are good (Wieczorek, 2023). However, due to the subjection of values to revision and the potential for unconsidered or unreflective uses of values, there exists a vulnerability where a value's propensity to outside influence might be exploited (Grenberg, 2007).

When someone has deliberated on how certain values might help them achieve an objective, practically speaking the only way to see if their judgment is correct is to test it through some means. One may believe their judgment to be correct to the point of feeling certain, but if they wish to produce some perceivable effect then belief is not enough, they must demonstrate. Pragmatism suggests two ways this might be possible (Dewey, 1922b). First, a person may act on their values in actuality, physically acting on the world whether that be with their hands, mouth, keyboard, etc. Second, a person may imagine what acting on their values might look like. In this second case, if one wishes to think on the impacts their actions may have in the real world, then their imagination should follow suit in the sense that they would do well to account for relevant features of real life. Imagination is useful for testing how a value contributes to a goal, particularly in scenarios where acting tangibly is not yet preferable. If a person has any desire to eventually enact this value in the real world, then they would do well to let their imagination include the relevant details of the real world as best they can. As a result, they will likely include details they think are physically possible, perhaps details they have experienced in the past, and responses and consequences they might reasonably guess based on these details. Dewey (1922a) calls this

process Dramatic Rehearsal and it along with physical enactments of judgment allow one to test out their values in the world.

Once an enactment of value, real or imagined, is carried out, it is available for reflection. In the pragmatic tradition, reflection is a conscious activity which aims at “suggesting alternatives that might otherwise be overlooked, and stimulating greater consistency in judgment” (Dewey and Tufts, 1936, p. 175). The aim is to examine a previous decision to see what worked and what did not work. Succeeding in one’s objective may help to confirm the relevant strategy while alternatives may arise out of instances of failure. If one’s enactment of value fails to produce the expected effect, then if they wish to continue their project, they would do well to consider what they may have missed. If a missed detail is discovered, the decision-maker can utilize this information to help suggest an alternative. For example, imagine some person values a certain friendship and attempts to enact this value with a loving embrace. However, the friend recoils from the embrace, denying the advance. Why might this attempt have failed? Perhaps there is some unresolved hurt that needs dealing with first, perhaps there is danger present and now is not the time, or perhaps the hug is simply unwelcome. Whatever the case, the failed attempt to enact a given value requires a readjustment in judgment should another attempt be made if deemed appropriate.

One option for improving judgment may be to adjust the values someone holds. Perhaps, to better fit, one’s attitudes need to be increased or decreased by some degree. Perhaps, if a value is to fit with a new circumstance, a redefinition is in order. As specific enactments of value fail in some measure to achieve the goals they aim for, and as new advancements in society suggest re-examinations of old enactments of value, new and possibly improved understandings of those values may be necessary. It is for this reason that values may be adjusted in practice. As a result, it is likely inappropriate to think of values as fixed entities to be found “already in existence” (Gouinlock, 1978, p. 223). Values are concepts which allow for reflection on past and present circumstances. However, since they help qualify what is desirable, pragmatism understands values to be predominantly future oriented concepts, even if someone’s desire is merely to maintain the current state of the object of their desire. As a result, values can be considered works in progress to the extent that they help guide ongoing decisions that may require adjustment upon reflection.

Values may also be considered social concepts. Humans are social creatures whose goals and actions bump into each other – aiding, avoiding, or hindering one another. Values are by extension often social concepts in two senses. First, values are social in the ways that individual value practices are often experienced by others (Wieczorek, 2023). Human decisions often impact the world around them which includes both nature and other creatures. Human actions have the power to profoundly change the conditions of their environment. If certain conditions are able to be perceived by others, then they may also be able to reflect on those conditions to inform their own decision-making. Thus, any change to the environment enacted in accordance with some value might be socially perceived and reflected on. Second, values can be social in that many individuals may have an agreed upon understanding of certain values. This latter social depiction of values pushes back on the notion that values could be universal. Some goods may very well be universal, but none will have a perfect judgment of such a good (Grenberg, 2007; Kaag, 2013). If ‘good’ is something human made, whether through one’s work or sentiment of desire, then judgment will be required in the making. If ‘good’ is an *a priori* concept to be found in the world, then judgment will be required in the finding. Explicit values, as concepts which require judgment, will thus likely require some agreement in judgment if an understanding is to be socially practiced.

The ability to perceive and experience some enactment of value is not limited to some explicit conceptualization of that value, it may be implicit as well. For example, De Monticelli (2019) thinks of how a ferocious mask might frighten a child without that child knowing the value of ferocity. That, she claims, “says something about the nature of values” (p. 118). I argue that this says that though the contents of some enactment of value is not limited to some predefined entity (Boenink and Kudina, 2020) values are not concept free either. A fierce mask may frighten a child; that fear and subsequent desire to perhaps get rid of the mask suggests some implicit concept of what it means to them. As soon as that child attempts to explain all of this, then explicit conceptions will be needed. Sometimes one discovers that the conceptions they are attempting to explicate have already been made by others who preceded them, and those earlier depictions may have an influence on the value holder’s understanding. What this example means to demonstrate is that values might be conceived of as works in progress which may arise implicitly and form through social discourse or practice.

When some value is adjusted as a result of a change in personal or societal circumstances it can be considered to be formed or take shape through such processes. Values, as discussed in section 1.2, may help guide people to make decisions in their life; they can serve as a part of

people's regular habits or practices. Through these practices a value's meaning might be given or found (Boenink & Kudina, 2020). In other words, to practice some value may allow someone to come to a deeper understanding of what it means to them, either implicitly or explicitly. However, a person's practices may need to change if they are to accommodate a relevant change in circumstances. If such a change in practice occurs so as to warrant the re-examination of some value then the personal or societal meanings of said value may have to change accordingly. In other words, "the core meaning of values is constantly being (re-)experienced and worked out within these practices" (Boenink & Kudina, 2020, p. 455).

To the extent that the meaning of values continuously takes shape within everyday interactions, there exists a vulnerability. As this section argued, no attitude towards a given value is final. Values are subject to change which may come about through societal influence (Wieczorek, 2023). Likewise, since people have the potential to unreflectively or inconsiderately use values, people may be vulnerable to this influence (Grenberg, 2007). If an agent were to be able to interject themselves into these everyday interactions, they might have the potential to change how values are enacted. Such an agent might suggest certain values as more or less important than they are currently understood in a given society or culture. They could capitalize on specific ways that humans fail, psychologically for instance, and trick people into adopting values they otherwise might not. To analyze this possibility, the next section will introduce and discuss manipulation.

Section 1.4. Manipulation may influence values

This section will discuss what manipulation is and how it functions before finally elaborating on how manipulation can tamper with a person's values. A common characterization in academic literature depicts a manipulator as a kind of puppeteer 'playing' with their targets by pulling at certain 'strings' to evoke specific responses. Such 'playing with' can go unnoticed and, in fact, may be quite a bit more effective the longer it does (Sunstein, 2016; Susser et al., 2019; Cohen, 2023). As a result, manipulation may inflict two potential harms to its victims: a loss of autonomy and the acceptance of unchosen values. Autonomy can be considered to be the ability to govern one's own life or to decide for oneself how they ought to live (Susser et al., 2019).

Before diving into the discussion on manipulation, the difference between deception and manipulation must be discussed as there is some disagreement over whether deception

encompasses manipulation or vice versa. This distinction is important as any two articles which discuss these as different concepts may, in fact, be concerned with the same thing. This will be particularly relevant in the next chapter which will dive into Dark Patterns, otherwise known as Deceptive Patterns. Despite the name, I will argue that Dark Patterns are predominantly concerned with the manipulation of end users, and so definitions here are warranted. The primary difference between deception and manipulation, I argue, is the outcome they are after.

Deception attempts to alter a person's belief about matters at hand such that they become mistaken or deluded (Cohen, 2023). There need not be a specific intention of what the deceived person is to do with that mistaken belief; the deception may merely conceal something the deceiver wishes to remain unnoticed. Deception, therefore, is inherently covert as once it is discovered, the mistake is revealed (Sunstein, 2016). For example, a pedestrian may lie that they have no spare change when asked for some by a person experiencing homelessness. The pedestrian may not have a specific outcome in mind aside from merely hiding the fact of their available wealth. However, deception may, sometimes, go as far as to seek a more specific outcome from the alteration of the deceived's belief (Susser et al., 2019). Here deception crosses over explicitly into the realm of manipulation.

Where deception aims to alter belief, manipulation seeks to alter a person's behavior, often against their self-determined choices. The goal of manipulation remains distinct from deception. Manipulation "interferes with the workings of judgment" i.e., the ability to reflect and deliberate, rather than altering the input of judgment as deception does (Cohen, 2023, p. 299). While both deception and manipulation therefore interfere with judgment in some way, there is a specific aim of what a manipulated person is to do, often in the favor of the manipulator. This is the feeling of puppeteering mentioned earlier as experienced when one attempts to govern a person's actions and decisions by pulling at certain 'strings' (Mildner et al., 2023). For example, perhaps an advertiser wishing to pull at the heartstrings of a viewer decides to evoke nostalgia by using images of happy families in their latest advertisement, implying that purchasing their product will make their customers likewise happy. While the degree of manipulation at play might be rather soft, it is at work in this example in the sense that the advertiser's aim is to alter the emotions of their audience to better guarantee the desired behavior, namely purchasing the advertiser's product. Conceivably, the audience may well be aware of this process, and so manipulation need not be covert (Sunstein, 2016).

Manipulation fails to respect the reflections and reasons of a decision-maker. In line with this notion that manipulation seeks to alter behavior, Sunstein (2016) argues that “an effort to influence people’s choices counts as manipulative to the extent that it does not sufficiently engage or appeal to their capacity for reflection and deliberation” (p. 216). Manipulation is problematic in the ways it seeks to overwrite a person’s decision-making processes insofar as a person’s decisions are based on the choices and reasons available to them. When successful, manipulation makes the manipulated person an instrument of the manipulator. As a person’s decisions fall in line with the manipulator’s desires their own autonomous deliberations may go unacknowledged and disrespected.

The intent of manipulation is to get the manipulated person to do as the manipulator pleases. These aims may sometimes be thought to be in the best interests of the manipulated, as is the case with some paternalistic justifications (Susser et al., 2019). However, manipulation carries a risk that the interests of the manipulator will be prioritized over the subject being manipulated. In either case, “the problem with the manipulator is that he lacks relevant knowledge about the chooser’s situation, tastes, and values” (Sunstein, 2016, p. 228). Though large tech companies like Google sometimes justify manipulative acts by claiming to be able to know their users better than they know themselves (Zuboff, 2015; König, 2022), they necessarily lack such first-hand knowledge. So, any conclusions about their user’s situations, tastes, and values that are not made explicit by their customers are, at best, inferences or presumptions. The experiences, tastes, and values a given manipulator can know best is their own. As a result, it is often easier for a manipulation to be aligned with the manipulator’s own interests rather than the interests of the manipulated.

Through the act of a successful manipulation, a manipulator may change the context through which the manipulated person acts. If someone were to be manipulated into allowing their personal data to be collected by some company then their context would change such that the company now has access to this data. Furthermore, this context carries the risk of reflecting the manipulator’s interests rather than those of the manipulated. This is acutely the case when manipulation is systematically enacted, i.e., sustained “through ongoing social feedback and comparison” (König, 2022, p. 1378). As a result, manipulation has the potential to not only change targeted behaviors, but also change an individual’s broader practices as they relate to the context of manipulation. Practices are developed within contexts and, like values, may have to adapt to new and unique circumstances (Dewey, 1922b). For example, say an individual is purchasing

their first smartphone and, when setting it up, is asked to consent to the phone using their personal data. They decline for personal reasons, but each day thereafter the phone nags the user with consistent pop-ups, re-asking for data. One day, tired of the inconvenience of the pop-ups, the user gives in and accepts the phone's use of their personal data (Luguri & Strahilevitz, 2021). This user now lives in a new world where their data is out there to be utilized as desired by, at least, their phone company; their context has changed and, as a result, their practices may need to adjust so as to account for this change.

Due to the potential for manipulation to effect a person's practices, it may also have an impact on personal values. On this account, the goal of manipulation is "to hurry, evade, or undermine deliberation, and thus to encourage decisions that may or may not align with an individual's deeper, reflective, self-chosen ends and values" (Susser et al., 2019, p. 9). This not only deprives the manipulated of their autonomy and the respect that entails (Sunstein, 2016), it also has the potential to influence how that individual chooses or changes their goals and values. Continuing with the new smartphone user example, this person now lives in a reality where their personal data is owned by a company to be used as desired. Perhaps this decision was made out of frustration and becomes something this person regrets. In reflecting on this decision and looking forward they have a few options. They may try to reverse their decision which, if possible, will return them to the manipulative nagging of before. They may try to account for this new reality and disable any targeted advertising (which often comes with consenting to the use of data, see Papadopoulos et al., 2017) and practice awareness of further manipulative strategies. Or they may try to justify their decision, convincing themselves that it was a 'good enough' choice to make. In these latter two cases, new values or changes to old ones will likely be required to accommodate these new practices. This person's value of privacy, for example, may need to adapt to include exceptions in the first case or decide that privacy is perhaps less valuable than they previously thought in the second case. This would not have been the case had the individual's initial autonomous decision been respected.

In short, when manipulation effects behavior it also effects practice, so to the extent that values take shape in practice manipulation may influence values. The changes that manipulation can make to one's context, practice, and values may have a broader influence on other behaviors related to those contexts, practices, or values. As discussed earlier, choices and decisions play a crucial role in value practices. When those decisions are not fully autonomous the consequences, including changes to values, may go unnoticed or unchallenged. Any values

adopted or adjusted as a result of manipulation may persist durably and dynamically across a multitude of situations. Values are deeply held and may be deliberately enacted, yet they are susceptible to other people who act with their own intentions as well. For example, in an effort to bypass reflective systems, manipulations may seek to foster habitual responses that steer towards the manipulator's intentions (Sunstein, 2016; Susser et al., 2019). Any values at play in the context of that habituation may also be informed by that context. Manipulation, therefore, may favor the values of the manipulator over the manipulated in persistent and pervasive ways. Now that the way manipulation can alter values has been argued, the next chapter will get into how this can happen online through the use of Dark Patterns.

Chapter 2 - Values and Dark Patterns

Chapter 2 Introduction

This chapter will explain what Dark Patterns are, where they come from, how they manipulate, and argue for why that matters concerning values. Dark Patterns are digital interfaces which “make it difficult for users to express their actual preferences, or manipulate users into taking certain actions” (Luguri & Strahilevitz, 2021, p. 44). Through the presentation of on-screen options, Dark Patterns trick, frustrate, confuse, or otherwise coerce end users into making choices that are often not in their best interest (Gray et al., 2018). Dark Patterns arose in User Experience (UX) design as the results of A/B testing performed rapidly and in real time online. An A/B test randomly issues two similar but notably different variants of a given interface to the users of a software application and measures which is more successful at achieving the designer's goals. Through the use of these tests within UX, developers found design patterns which engineered seamless use but also seamless influence (Luguri & Strahilevitz, 2021; Mildner et al., 2023).

A point of discussion gaining traction in UX and HCI today is how these designs often get end users to act against their own self-interests (Gray et al., 2018; Susser et al., 2019). Despite the evidence of their manipulateness, Dark Patterns are liberally used, especially on the most popular websites and applications (or apps) available today. Though end users often struggle to identify specific implementations of Dark Patterns, many negative impacts are personally experienced in ways that leave them feeling helpless (Luguri and Strahilevitz, 2021; Bongard-Blanchy et al., 2021; Gray et al., 2021). I argue that, based on the evidence, this is deliberate.

Platforms which implement Dark Patterns know of their effectiveness and (should) know of the harm they can cause. Social media platforms, or 'platforms' for short, are online virtual spaces that are engaged through a website or app where end users can socially connect to create, share, consume, and engage with content.

Due to the funding strategy of platforms, namely by Venture Capital, the designers' fiduciary duties to their funders may be held in higher regard than their responsibility to respect the autonomy of their end users. This results in a system which consistently holds a platform's values as higher than personal or societal values and attempts to manipulate end users to do the same. As a result, to the extent that values are formed in practice, one's personal values are likely to be unduly influenced by such a system.

Section 2.1. A brief history of Dark Patterns

Dark Patterns are reusable software solutions which exploit a user's decision-making vulnerabilities (often psychological) through the way they organize information on screen. Specifically, Dark Patterns organize on screen information in ways that are considered to exploit cognitive biases. They trick, or coerce the user into taking specific actions that are often against their own self-interest (Gray et al., 2018; Susser et al., 2019; Mildner et al., 2023). Dark patterns were first discussed in the field of User Experience (UX) design by Harry Brignull in 2010 as "tricks used in websites and apps that make you do things that you didn't mean to, like buying or signing up for something" (Brignull, 2010). Such patterns were often integrated into a platform's algorithm which dictated proprietary methods of delivering content and were thus protected as a trade secret (Zuboff, 2015; Luguri & Strahilevitz, 2021). Since then, taxonomies have been developed which categorize the increasing corpus of Dark Pattern types and the design strategies to which they contribute.

As of 2023, Mildner et al. has composed a list of at least 81 specific Dark Patterns (p. 7). Other researchers like Luguri & Strahilevitz (2021) believe such a list may be too long and choose to limit their list to a more modest 27 variants (p. 53, see Table 1 at the end of this section). Some of the most common variants include 'nagging' which repeatedly prompts the user to accept something the platform desires, sometimes without the option to permanently decline. Google has used a nagging approach to their location service agreements where an end user can 'Agree' or 'Disagree' to allow Google to collect location data on a user's device at all times (Gray et al.,

2018). If the user disagrees Google will ask again at another time, but if the user agrees Google will stop asking and continuously collect location data. Another common example is termed 'privacy zuckering' which tricks end users into sharing more information about themselves than they meant to (Gray et al., 2018). Named after Mark Zuckerberg, this dark pattern has been infamously utilized by Facebook where they collected and used more personal user data than their users could reasonably determine from Facebook's terms of use.

Each Dark Pattern follows in a similar pattern. Through the presentation of on screen options Dark Patterns make certain options easy to choose and other options difficult or impossible to choose where the easier options tend to be ones that benefit the platform most. A potential reason for the vast number of dark pattern variants is simply "because the internal, proprietary research [performed by platforms] suggests Dark Patterns generate profits for the firms that employ them" (Luguri & Strahilevitz, 2021, p. 45). In other words, designers are good at finding many specific and profitable ways of influencing the decisions of their users. Through their influence on behavior, Dark Patterns manipulate end users, interfering with their capacities to autonomously reflect and deliberate on relevant decisions in noticeable ways (Bongard-Blanchy et al., 2021; Gray et al., 2021). How Dark Patterns accomplish this and the ways they interfere with end users will be explained in this chapter.

The origin of Dark Patterns traces back to early internet commoditization in multi-channel markets. A channel, here, is a route or method for a given firm to "market their products and services" (Langley and Leyshon, 2017, p. 17). Prior to the advent of the internet, markets relied more on physicality - whether that be a location where exchange takes place, such as a stock exchange building, or a physical product. Buyers and sellers were connected via businesses which intermediated that connection linking their value to each other more or less directly. What the internet changed, according to Langley and Leyshon (2017) was the role of businesses in this intermediation. Online spaces were no longer merely a place that fostered connections between buyers and sellers, they became platforms "actively curating connectivity" (p. 15). These online spaces were not only connecting people together, but fostering creative interactions between users in the form of content. As content creators, users are considered to co-create value in the sense that they create reasons for other users to use the respective online space (Langley and Leyshon, 2017).

Many platforms only became profitable businesses well after their creation due to a popular modern investment model that allowed for *ex post* financial justification known as venture capital. Venture capital investments are inherently risky as they fund “early-stage emerging companies with little to no operating history but significant potential for growth” (Baldrige, 2022). Thus, though emerging platforms may not have any business experience, as long as they have a plan to scale up within their respective market, they retain the potential for a huge return on investment. Venture capital firms know of the risky nature of their investment and so expect losses. However, they also expect that the platforms which do succeed will be ‘home runs’ and realize returns that greatly exceed the losses on the failed platforms (Langley & Leyshon, 2017). Thus, there is a high expectation for platforms to prove that they are one of those home runs. Assuming that a venture capital firm has funded enough of these high-risk but high-reward startups, the firm is considered likely to gain more than they lost. As a result, venture capitalism became the machine that drove platform startups which required “confidentiality and privacy to protect intellectual assets” (Fields et al., 2020, p. 463). Since a firm’s portfolios of the platforms that they invest in are sensitive, so too is each platform’s algorithms, and the whole package is protected as trade secrets (Luguri & Strahilevitz, 2021). This may be the reason that platforms have historically received little regulatory oversight and, as a result, may not have been incentivized to evaluate any problematic outcomes of their connectivity algorithms (Zuboff, 2015).

The term ‘algorithm’ as it is used colloquially often denotes a computational procedure that is “distributed, probabilistic, secret, continuously upgraded, and corporately produced” (Seaver, 2017, p. 3). Distributed, here, refers to the ways algorithms are disseminated throughout the societies they enter. As such, algorithms become culturally responsive in the sense that their development often includes multiple regional variants which make adjustments based on the relevant features of various local cultures. As platforms collect vast amounts of digital data they use algorithms as one of the main tools to automatically analyze this data to figure out ways to keep a variety of people coming back. As platforms are digital spaces, ‘coming back’ has to mean something different to a traditional brick-and-mortar store. As such, a platform and its algorithm measures success, in part, with attention, i.e., the number of users and the amount of time they spend using the platform (Langley & Leyshon, 2017; König, 2022).

A platform’s algorithm often measures the ability to successfully capture end-user attention through the use of rapid A/B testing, otherwise known as continuous experimentation. In an A/B test, “two software variants, denoted as variant A and variant B, are compared by evaluating the

merit of the variants through exposure to the end-users of the system” (Quin et al., 2024, p. 2). These two variants are pushed out to random users of a given platform, some experience variant A while an equal portion of others experience variant B. The software variant that performs better for some specified purpose is either selected as a feature to appear in a broader design framework and/or pitted against other variants in further A/B tests. Furthermore, this process can be, and often is, automated (Quin et al., 2024). This, I argue, is what the term ‘algorithm’ often means in popular western culture: the automated continuous experimentation that leads to the personalization of apps and websites. Since a major concern for platforms is keeping end user attention on their website or app, these A/B tests were initially optimized for attracting and maintaining attention. Such optimization came about from the recognition that platforms need to “invest in behavior design to create new habits that repeatedly bring producers and consumers back to the platform” (Choudary, 2015, p. 21).

Once a platform has proven its capacity to capture and keep an individual’s attention, advertising becomes a lucrative revenue model. Software companies quickly realized that advertisers may purchase a selection of screen time and digital real estate on a given platform to capitalize on the attention the platform had garnered. However, this softer method of behavior design has today evolved into what is known as behavioral-targeted advertising (Zuboff, 2015; Langley and Leyshon, 2017). These types of advertisements offer ultra-specific products based on an end user's data generated from platforms and sold in Real Time Bidding (RTB) auctions to advertisers. As a result, this system that posited consumers as users which co-create value has shifted towards treating these users as products themselves (Papadopoulos et al., 2017).

A platform’s continuous experimentation has evolved to the point where, today, their behavioral influence extends far beyond mere attention and advertising. Through the use of A/B tests, firms can find very subtle ways of influencing many different online behaviors of their end users (Luguri & Strahilevitz, 2021). Attention still remains an important behavior to influence, but other behaviors have been discovered to be influenceable as well such as the acceptance of cookies, the purchase of unwanted products or services, and the giving up of personal data (Gray et al., 2018; Mildner et al., 2023). No longer are the platforms which utilize these Dark Patterns satisfied with keeping consumers coming back to them, they now seek to directly influence many specific behaviors of their end users.

Dark patterns are a product of this so-called 'platform capitalism', or the systematized financing of digital platforms to "realize a return on capital" (Langley & Leyshon, 2017, p. 22). Table 1 provides a summary of the numerous variants of Dark Patterns and briefly describes what they do (for screenshots of some examples, see Appendix A). The story of their origin is important to see how systemically embedded they are. They are not just a feature of the last decade, but something that has been evolving over the past few decades at least. Dark patterns are intertwined with novel investment strategies, the invention of the internet, and other technological advancements, as well as behavioral psychology among other things (Langley & Leyshon, 2017; Thies et al., 2019; Luguri & Strahilevitz, 2021). The history of Dark Patterns is important to know because what is wrong with Dark Patterns, as argued in the next section, stems from the system they were born out of, namely one that has learned that the most profitable ventures are those that prey on people's weaknesses and vulnerabilities (Zuboff, 2015; Gray et al., 2021). This may not be surprising, but this thesis will argue why it is especially concerning when it comes to Dark Patterns.

Category	Variant	Description	Source
Nagging		Repeated requests to do something the firm prefers	Gray et al. (2018)
Social proof	Activity messages	False/misleading Notice that others are purchasing, contributing	Mathur et al. (2019)
	Testimonials	False/misleading positive statements from customers	Mathur et al. (2019)
Obstruction	Roach motel	Asymmetry between signing up and canceling	Gray et al. (2018), Mathur et al. (2019)
	Price comparison prevention	Frustrates comparison shopping	Brignull (2020), Gray et al. (2018), Mathur et al. (2019)
	Intermediate currency	Purchases in virtual currency to obscure cost	Brignull (2020)
	Immortal accounts	Account and consumer info cannot be deleted	Bösch et al. (2016)
Sneaking	Sneak into basket	Item consumer did not add is in cart	Brignull (2020), Gray et al. (2018), Mathur et al. (2019)
	Hidden costs	Costs obscured/disclosed late in transaction	Brignull (2020), Gray et al. (2018), Mathur et al. (2019)
	Hidden subscription/forced continuity	Unanticipated/undesired automatic renewal	Brignull (2020), Gray et al. (2018), Mathur et al. (2019)
	Bait and switch	Customer sold something other than what's originally advertised	Gray et al. (2018)
Interface interference	Hidden information/aesthetic manipulation	Important information visually obscured	Gray et al. (2018)
	Preselection	Firm-friendly default is preselected	Bösch et al. (2016), Gray et al. (2018)
	Toying with emotion	Emotionally manipulative framing	Gray et al. (2018)
	False hierarchy/pressured selling	Manipulation to select more expensive version	Gray et al. (2018), Mathur et al. (2019)
	Trick questions	Intentional or obvious ambiguity	Gray et al. (2018), Mathur et al. (2019)
	Disguised ad	Consumer induced to click on something that isn't apparent ad	Brignull (2020), Gray et al. (2018)
	Confirmshaming	Choice framed in a way that makes it seem dishonorable, stupid	Brignull (2020), Mathur et al. (2019)
	Cuteness	Consumers likely to trust attractive robot	Cherie & Catherine (2019)
Forced action	Friend spam/social pyramid/address book leeching	Manipulative extraction of information about other users	Brignull (2020), Bösch et al. (2016), Gray et al. (2018)
	Privacy Zuckering	Consumers tricked into sharing personal info	Brignull (2020), Bösch et al. (2016), Gray et al. (2018)
	Gamification	Features earned through repeated use	Gray et al. (2018)
	Forced Registration	Consumer tricked into thinking registration necessary	Bösch et al. (2016)
Scarcity	Low stock message	Consumer informed of limited quantities	Mathur et al. (2019)
	High demand message	Consumer informed others are buying remaining stock	Mathur et al. (2019)
Urgency	Countdown timer	Opportunity ends soon with blatant visual cue	Mathur et al. (2019)
	Limited time message	Opportunity ends soon	Mathur et al. (2019)

Table 1. Summary of existing Dark Patterns, taken from Luguri & Strahilevitz (2021, p. 53)

Section 2.2. How Dark Patterns manipulate

The paramount desired behaviors a platform has for its users is either to get that user to pay for something on the platform, to get the user to do something someone else has paid for, or to nudge a user in either of those directions (Papadopoulos et al., 2017; Langley and Leyshon, 2017). A major goal of platforms, based on the venture funded laissez-faire system, is of course to make money. To be profitable, the behavior design of platforms generally must find a way to get their end users to pay for something (Gray et al., 2021; Luguri and Strahilevitz, 2021). Dark Patterns are an effective way of manipulatively achieving this goal. This section argues that Dark Patterns manipulate people by exploiting cognitive biases in the design of the on-screen choices (Willis, 2020). These cognitive biases encourage the user to act automatically which undermines any careful or considered thinking – a hallmark of manipulation (Sunstein, 2016; Susser et al., 2019). They achieve this form of manipulation through continual experimentation and design iteration which tests and implements ways which have been proven the most successful in getting the end user to do as the platform wishes (Quin et al., 2024). These methods are not only effective but often negatively experienced, in some capacity, by end users as they coerced into decisions that they would otherwise not wish to make (Gray et al., 2021). This section starts with a technical example of how manipulation can be automatically formed in design.

The A/B testing utilized for capturing and keeping a user's attention is now being used to design for optimally effective manipulations. For example, the United States Federal Trade Commission (FTC) recently took action against the personal finance company Credit Karma for false advertising (Division of Financial Practices, 2022). Through automated A/B testing, Credit Karma learned that they could increase click-through rates on their advertisements with a false statement. Falsely stating that their users had been “‘pre-approved’ for a credit card” yielded more optimal click-through rates than the same advertisement stating viewers had “‘Excellent’ odds of being approved” (Division of Financial Practices, 2022, p. 3). As a result, Credit Karma chose to deploy the more deceitful yet effective option. Although neither the attempt to find the most effective advertisement nor the use of deceit in advertisement are new, what is new is that these processes are now automatically tested in real-time and in an online environment at the fingertips of the customer. This continuous experimentation efficiently discovers effective ways of getting users to do what the designers want (Willis, 2020).

In terms of their effectiveness, Luguri and Strahilevitz (2021) claim to have performed the first studies that publicly demonstrate the power of Dark Patterns. Prior to this point, the evidence of the effectiveness of Dark Patterns was based predominantly on their ubiquity. After Harry Brignull's coining of the term in 2010, researchers, practitioners, journalists, and policy makers all began to identify a wide corpus of Dark Patterns found in the many corners of the internet (Gray et al., 2018; Susser et al., 2019). This led Luguri and Strahilevitz (2021) to claim that the prevalence of Dark Patterns suggests their effectiveness and to attempt to prove that claim. Their own experiments tested known Dark Patterns (e.g., false hierarchy and confirmshaming, see Appendix A for examples of each), which ranged from mild to aggressive, against no Dark Patterns (i.e., a straightforward choice). The results of these experiments suggest that "Dark Patterns are strikingly effective in getting consumers to do what they would [otherwise] not do when confronted with more neutral user interfaces" (p. 46). Dark patterns are so effective, in fact, that they may increase the acceptance of platform preferences by two to four times over the neutral interface rates (Luguri and Strahilevitz, 2021).

The way Dark Patterns achieve this level of effectiveness is by interfering with a user's reflective capacities and judgment. They undermine the user's deliberation and encourage decisions which have no regard for that individual's intentions or values (Susser et al., 2019). In psychological terms, Dark Patterns are considered to encourage what is known as 'System 1 decision-making' and discourage 'System 2 decision-making'. "System 1 is the automatic, intuitive cognitive system, prone to biases and to the use of heuristics, while System 2 is more deliberative, calculative, and reflective" (Sunstein, 2016, p. 222). This 'encouragement' is performed by "altering online choice architecture in ways that are designed to thwart users' preferences" and exploit their cognitive biases (Luguri and Strahilevitz, 2021, p. 52). For example, 'roach motel' is a dark pattern which makes it easy to sign up for an online service and much more difficult to cancel said service. Like roaches, such subscriptions can be difficult to eliminate. The tactic depends on making a user more comfortable with signing up due to the promise that they can cancel anytime, while actual cancellation is made confusing and tedious. Sony has been accused of using a roach motel in their PlayStation+ subscription (Luguri and Strahilevitz, 2021; see Appendix A for a screenshot). Through the exploitation of cognitive biases by the curation of on screen choices, Dark Patterns undermine a user's reflection and deliberation.

As manipulation "does not sufficiently engage or appeal to [one's] capacity for reflection and deliberation" Dark Patterns may therefore be understood to be manipulative (Sunstein, 2016, p.

216). They “interfere with the workings of judgment” (Cohen, 2023, p. 299), pulling on different psychological strings to influence the decisions borne out of that judgment. The coercive strategies that Dark Patterns employ consistently prioritize action possibilities that shareholders, say, would find desirable. When using an interface that utilizes Dark Patterns, a user’s potential actions are limited in the ways Dark Patterns frustrate their available actions which do not benefit the platform. This frustration may materialize in ways that detract from options users should (e.g., legally) have within the interface but are undesirable for the company to let their users take (Gray et al., 2018). For example, putting an item in someone’s virtual basket while online shopping is arguably something the end user should have control over. However, the Dark Pattern known as ‘sneak into basket’ (see Table 1), attempts to covertly add an item to the shopper’s basket which they may unknowingly pay for if the item remains hidden. Such tactics can be considered to be anti-user in the sense that they consistently deprioritize the judgment of users.

These manipulated decisions may be ones that an end user would rather not make had they been able to determine their own choices (Gray et al., 2021). Though certain options may be the most desirable to a given user, when they are made to be more difficult to pick than others, people may reasonably choose options of lesser resistance. For instance, though an end user may prefer to deny consent to the use of cookies, if that option is obfuscated or made more frustrating and the option to consent is made less difficult, then they may be persuaded to consent despite their preference. This is supported in recent research, both in terms of effectiveness and in consumer self-reporting. When users are made aware of Dark Patterns, they may regret or otherwise dislike the ultimate decision they made as the result of Dark Patterns, indicating a mismatch between the decision made and their preferences (Luguri and Strahilevitz, 2021; Bongard-Blanchy et al., 2021; Westin & Chiasson, 2021; Borberg et al., 2022). For example, many people are dependent on online services that use cookies which track their personal data across websites and apps. However, despite feeling concern over consenting to the use of cookies or outright distrusting websites that use cookies, users may feel that they have no choice but to consent due to Dark Patterns like nagging and privacy zuckering (covered in section 2.1.). This may result in a “perceived lack of control over their data” (Borberg et al., 2022, p. 10). It is this kind of disregard for people’s self-determination through the undermining of their deliberation in the interest of a platform that makes Dark Patterns manipulative (Susser et al., 2019).

Though Dark Patterns may have specific objectives, end users often lack the ability to comprehensively understand what those objectives might be. In their recent research, Gray et al.

(2021) ran a survey seeking to understand how users experience the manipulative power of Dark Patterns. Their findings suggest that people struggled with initially identifying Dark Patterns as well as any sense of manipulation that followed that initial judgment. This covert aspect of Dark Patterns can make personal inquiry into meaningful options difficult (Tschaepe, 2018). However, many study participants were widely capable of noticing that “something is ‘off’” although unable to describe precisely what may be driving that feeling (Gray et al., 2021, p. 19). When a participant was able to identify a specific instance of manipulation, they were likely to describe it as distressing or confusing. Gray et al. (2021) argue that this kind of research can help raise awareness. However, though important, awareness may not be enough to counteract the “likelihood to be influenced by manipulative designs” (Bongard-Blanchy et al., 2021, p. 771). To really counteract Dark Patterns, more than mere awareness may be required, which this thesis will discuss in chapter 3.

So, Dark Patterns are effective at manipulating users of platforms who struggle to discern what is happening to them. This alone is concerning but the pervasiveness of Dark Patterns may indicate a more sinister issue. For more than a decade now, platforms have been using Dark Patterns in online social spaces, shaping the habits of consumers for lasting engagement. Together, venture capitalists, advertisers, and platforms have formed an architecture for e-commerce that has the potential to “sustain practices that heavily intervene into society and social relations on a structural level” (König, 2022, p. 1378). Though it is individuals who experience the negative repercussions of Dark Patterns, any one individual is not so important for a successful platform so long as the systemic structure remains intact and broadly effective. Thus, end users become an expendable commodity (Zuboff, 2015). In other words, Dark Patterns stem from a system that values its people predominantly as means to a profitable end. So, what happens to the people living in such a system that constantly seeks to manipulate their behavior? Before answering that question in section 2.4 and in the subsequent chapter it will be necessary to clearly demonstrate that people do live in such a system.

Section 2.3. The systemic nature of Dark Patterns

Now that the ways Dark Patterns can manipulate has been explained, this section will detail how Dark Patterns exist as a part of a systemic issue which is, namely, the acceptance of rapidly performed mass manipulation. Here I mean acceptance in a descriptive sense, not a normative one. I do not mean to say that Dark Patterns are or should be accepted in a legal, regulatory, or

standard sense but that Dark Patterns have become so prevalent in online platforms that many of their instantiations have come to be expected (König, 2022). This section will argue that many times Dark Patterns are used in tandem with one another to comprise a system that is more complex and potentially more difficult to resist in terms of their manipulative power (Westin & Chiasson, 2021). This systemic embedding of Dark Patterns within platforms used in everyday interactions have become normalized and it is something that both the industry and end users may have come to expect. Part of this expectation may be a financial one, where executives and designers may feel a stronger sense of responsibility towards their investors than their end users. This financial responsibility may be seen to justify the use of Dark Patterns either implicitly or explicitly in the development context (Gray and Boling, 2016). The end result of this system, I argue, is that the end user regularly experiences manipulation which degrades their consumer control (Luguri and Strahilevitz, 2021).

I argue that Dark Patterns comprise a systemic problem for the end users of platforms. An issue is systemic when it relates to the whole of some system or can be found distributed throughout it. It need not be found in every corner of the system, but cannot be isolated to one or two instances within the system either. To be systemic, something must be found throughout a system such that a relevant person who participates in that system is likely to experience some aspect of that thing. A systemic problem is a problem found throughout and within many different aspects of a system. This section will demonstrate how Dark Patterns are distributed throughout the system of online social media platforms. The impact of Dark Patterns, which the previous section argued is often problematic, is experienced at many different points throughout a platform capitalistic society.

Dark patterns are numerous and can work cooperatively to achieve some manipulative aim. The sheer number and variety of Dark Patterns has been offered as evidence of their effectiveness. As discussed in this chapter, some academic sources presume that the reason they are so numerous is because the industry firmly believes they work due to the evidence of A/B testing (Gray et al., 2018; Luguri & Strahilevitz, 2021; Bongard-Blanchy et al., 2021; Westin & Chiasson, 2021). Luguri & Strahilevitz's (2021) study further suggests that multiple Dark Patterns working together may increase the effectiveness of their manipulative strategy. In one of the experiments they ran, the researchers found that when a subject resisted a single instance of a dark pattern's coerciveness, they sometimes fell for another dark pattern implemented soon after the first. This suggests that even when people are capable of resisting some Dark Patterns, they may still remain susceptible to others. Dark Patterns may have a cumulative power that capitalizes on that

susceptibility, working together as a set of traps (Luguri & Strahilevitz, 2021). Furthermore, the strategy of some Dark Patterns relies upon their repeated use, increasing the number of times they may appear on a given platform. Such is the case for patterns like ‘nagging’ (described in section 2.1.) and ‘infinite scrolling’ where content endlessly generates as an end user scrolls down a page as implemented on platforms like Reddit and TikTok (Gray et al., 2018; Bongard-Blanchy et al., 2021, see Table 1 above and Table 3 in Appendix B). The liberal use of Dark Patterns on platforms indicates they may be a regular encounter for many end users.

Online platforms, and by extension the Dark Patterns they implement, are ubiquitous and pervasive to the point that they have become an everyday experience for their users (Langley and Leyshon, 2017; Westin and Chiasson, 2021; Bongard-Blanchy et al., 2021). Platforms are quite commonly used globally, but especially so in the Western world (Langley and Leyshon, 2017). Many users have come to depend on platforms for many different purposes such as social connection, navigation, or entertainment (Gray et al., 2021; Westin and Chiasson, 2021). Dark patterns may be present in as much as 95% of the most popular platforms (Bongard-Blanchy et al., 2021; Mildner et al., 2023). As a result, end users of platforms are likely to regularly experience Dark Patterns and may even expect them when navigating through platforms, even if they cannot identify specific instances of them. Research by Bongard-Blanchy et al. (2021), Westin and Chiasson (2021), and Borberg et al. (2022) have all pointed to attitudes of end users who experience quite a bit of frustration and distress from Dark Patterns but also a sense of resignation in some cases. Despite potentially disliking some platform, an end user may rely on that platform to fulfill certain wants or needs. Many participants in Gray et al’s (2021) research felt the use of specific platforms was necessary either to connect with people, as on social media, or accomplish a specific task, such as commuting through a ridesharing platform. When made aware of the manipulative efforts being made, these participants expressed sentiments of powerlessness as, despite these efforts, they still felt the use of a platform was necessary.

Dark patterns have become common software solutions in the platform industry which may be explained, in part, by the notion that a platform’s employees may feel a greater sense of responsibility towards their financial investors than their end users. Designers as guarantors is a useful concept to trace this problem through. The basic idea is that designers are, at least to some extent, responsible for the things they design. To this end, there is a recognized “need for licensure and accreditation in engineering disciplines” to ensure quality and protect public welfare (Gray and Boling, 2016, p. 979; Jang and Yu, 2008). Despite such accreditation processes

mandating ethical training, perhaps as little as 30% of design/engineering students are considered adequately prepared to handle ethical decision making, suggesting that the ethics training may be lacking. This gives rise to the question, what are these designers being trained to be responsible for and to whom? Gray et al. (2018) suggest that the industry expectations may be such that designers often focus on guaranteeing short-term successes, such as a successful launch, over long term ones, such as consumer well-being. The immediate financial health of a platform and the satisfaction of their investors may be aligned with this more short-sighted depiction of success. This may be further solidified by the expectation for startups to “demonstrate their capacity for revenue growth and thus cost-recovery to investors” (Langley and Leyshon, 2017, p. 24). Any ethical concerns beyond what can be guaranteed short-term might therefore fall to the wayside which may very well include the ethical concerns generated by Dark Patterns.

To summarize, Dark Patterns are numerous, distributed throughout the system of online platforms, and may be used conjunctively towards some manipulative aim by the platforms that utilize them. They are also commonly experienced by normal everyday users who have come to depend on platforms for various reasons. Finally, the focus of designers may be near-sighted such that long-term ethical concerns fall through the cracks of their licensure or accreditation training. In other words, designers may have a difficult time accounting for certain lasting problems within the system they help develop. The culmination of these factors constitute a fairly robust systemic problem. The ubiquity, normalization, and indifference regarding the manipulative nature of Dark Patterns compiles a bad deal for the end user who may systematically experience manipulation in such a way that it degrades their self-determination or self-control (Luguri and Strahilevitz, 2021). How this might happen in practice and the relevance of values is where this chapter turns to next.

Section 2.4. Dark Patterns may influence a person’s values

This final section of this chapter will argue for why and how Dark Patterns may influence personal values. As the last chapter argued, manipulation ‘plays with’ people in the sense that it pulls at different psycho-emotional ‘strings’ to steer a person’s decisions in the direction the manipulator desires. In the pulling of these strings, a person being manipulated has their judgment interfered which may decrease their ability to be self-directed while increasing their potential to be directed by the manipulator (Cohen, 2023). This evokes a sense that the manipulator may do with others as they please. For those of us who have experienced manipulation, “it presumes to decide for

us how and why we ought to live” (Susser et al., 2019, p. 4). Since how one ought to live may be guided by their personal values and Dark Patterns seek to influence how people live then Dark Patterns may likewise influence values. Several studies suggest this may be how end users experience Dark Patterns (Bongard-Blanchy et al., 2021; Gray et al., 2021; Westin & Chiasson, 2021; Borberg et al., 2022).

Since end users must deal with being ‘played with’ in daily online practices, it is reasonable to consider how and in what ways their values may be likewise ‘played with.’ Dark Patterns can be considered to establish or reinforce impulsive decision-making or make considered decision-making more difficult (Luguri and Strahilevitz, 2021). Thus, they may often emphasize short-term goals, both positively and negatively. For instance, the dark pattern dubbed ‘Testimonials’ (see Table 1) utilizes false or misleading statements about some product in the attempt to convince a customer to purchase or use said product. A common variant of this dark pattern comes in the form of fake reviews which have been known to plague online shopping platforms like Amazon (Hill, 2024). A store hosted on a shopping platform might utilize fake reviews to give the impression that their product is more reliable or trustworthy than it actually is, potentially convincing customers who value trust and reliability to purchase said product. In ways such as this, Dark Patterns can play with a customer’s values to try and trick them into believing the product aligns with their values.

Dark patterns may today be found as a more regular part of an end-users’ daily practice (Susser et al., 2019). Technological advancements have contributed to this in the ways they have made connecting online easier via technologies like mobile data and Wi-Fi. Furthermore, computational advancements have made mobile devices cheaper, more accessible, and increasingly useful. However, Dark Patterns themselves may have also contributed to this marked increase in online connectivity, which makes sense given the original purpose they served of capturing attention. For example, in recent years many young people have been found to spend around 5 and 1/2 hours per day on social media platforms, “almost a third of the daily active hours of any person” (Giraldo-Luque et al., 2020, p. 2). Such a high number of hours spent online is sometimes attributed to Dark Patterns which utilize addictive design elements including ‘on-by-default’ notifications as well as “pull-to-refresh, infinite scrolling, and auto playing media” (Mildner et al., 2023, p. 8). The end result is the regular and potentially adverse use of platforms within societies where platforms are prevalent.

As the previous chapter argued, values can be formed in practice; since Dark Patterns seek to change a person's behavior, Dark Patterns may reasonably be expected to have an influence on the values formed within online practices. As manipulation can lead a person "to act toward ends they haven't chosen... for reasons not authentically their own" (Susser et al., 2019, p. 9), they may be acting towards 'ends-in-view', or values, likewise not their own (Boenink & Kudina, 2020). Put another way, if a given personal value is not aligned with a particular attitude prompted by a dark pattern, and that dark pattern is successful in manipulating a given person's action, then that personal value may be considered neglected or overridden. Automatic habit formation, characteristic of System 1 cognition, may further de-emphasize certain personal values and promote values consistent with impulsivity. In cognitive psychology, impulsivity has been "associated with increased automaticity of behavior" (Ersche et al., 2019, p. 8).

When a Dark Pattern promotes automatic System 1 type responses, researchers like Giraldo-Luque et al. (2020) argue that they may suffuse a user's experience with impulsive and more self-centered values such as popularity or prestige. For example, when using social media that encourages personal photo and video sharing, like Facebook or Instagram, Westin and Chiasson (2021) argue that Dark Patterns can pressure users into feeling like they must post to validate their experiences. This goes so far in some cases where certain users feel they must post to prove they have done something worth remembering. Users of social media can be left feeling like a personal experience is not valuable if it is kept private or that it may need to be seen and liked by others to be considered valuable. Likewise, the more popular a post is, the more valuable the personal experience that the post is about may seem to such users. This idea will be revisited in more detail in chapter 3.

Dark patterns may achieve this influence by exploiting psychological impulses or biases (discussed in section 2.2) on a systemic scale. Values may constitute reasons for acting. So, when Dark Patterns prioritize System 1 cognition over System 2, values that align with impulsive System 1 behaviors may be emphasized over values that help govern deliberate System 2 behavior (Luguri & Strahilevitz, 2021; Sagiv & Roccas, 2021). Thus, values which fit better with impulsivity may be more likely to grab the attention of the person being manipulated by Dark Patterns. As the previous section argued, Dark Patterns are systemically embedded in online platforms. To the extent that a person spends time on platforms they can expect to get regular exposure to Dark Patterns and, in turn, many chances for impulsive values to grab their attention. Furthermore, the impulsive nature of Dark Patterns means they do not encourage users to be

reflective or deliberative which may leave their more carefully considered values potentially ignored, altered, or devalued. Thus, Dark Patterns systematically encourage decisions which deprioritize an individual's deeply reflective values (Susser et al., 2019). If this assessment is correct, and Dark Patterns do constitute a systemic manipulative practice, then platform capitalism signals somewhat of a societal degradation of consumer self-control and autonomy.

In the attempt to capture attention and govern online behavior, platforms create a space where it can be difficult to pay attention to anything else which may, in turn, make it more difficult to practice self-control. Psychologist and professor Dr. Gloria Mark has been running a study since 2004 on attention. Together with her team, Dr. Mark's results suggest that attention spans, or the ability to concentrate, has been steadily decreasing for professionals who use a computer screen. A potential reason the team explores is the increase in notifications in recent years which demand a user's attention (Smith, 2023). Notifications meant to capture one's attention as a part of a platform's alert system can detract from the object of one's focus, including other platforms. The many different notifications for the various platforms one may have access to on their device may consequently distract the user as they compete for the user's attention. Prolonged distraction may in turn lead "to problems of self-control and self-discipline" (Giraldo-Luque et al., 2020, p. 7). As a result, Dark Patterns may not only be manipulating specific, individual acts, but supporting a new norm where concentration and self-control are more difficult to maintain. Concentration and self-control are not only desirable characteristics, but essential disciplines for reflection and deliberation.

Resisting Dark Patterns and maintaining one's self-control and personal values can be a challenging prospect. The choice architecture implemented and the exploitation of cognitive biases attempt to make the choice the platform prefers much easier to select than any alternative options (Luguri and Strahilevitz, 2021). When personal values come into conflict with platform values, the resolution to that conflict may require prioritizing, balancing, or redefining the values at play (Boenink and Kudina, 2020). However, if an end user is consistently manipulated into defaulting to what the platform wants, it may likewise be much easier to default to the values of the platform rather than going through the effort of prioritizing, balancing, or redefining. These latter strategies are slower, more considered, and cognitively taxing when compared to automatically defaulting to the platform (Frankish, 2010). To rise to the challenge of resolving value conflict likely requires adequate motivation and effort by the value holder. To the extent that

a user has already defaulted to a platform's values, it may be that much more difficult to maintain certain personal values which conflict with those of the platform.

Dark Patterns may contribute to an undesirable value formation since they play a substantial role in the online experience and thus may influence what one believes their peers and society values. The harms to autonomy that Dark Patterns can cause may result in expressions of values that are unchosen and undesirable but difficult to resist. To the extent that values are formed in practice and serve to further guide action (discussed in chapter 1) Dark Patterns may promote the propagation of a platform's values through their use in practice. With recent studies demonstrating that young people already spend up to a third of their day on platforms (Giraldo-Luque et al., 2020), there is a legitimate space for concern that they will learn more about values from these platforms than in their offline lives. If these platform values are more self-centered, as argued earlier, then Dark Patterns may establish platforms as spaces which oppose communal values and make living together more difficult (Gouinlock, 1978). Some researchers argue that the manipulateness of Dark Patterns alone justify their regulation (Bongard-Blanchy et al., 2021; Luguri & Strahilevitz, 2021; Borberg et al., 2022). However, if this further concern is fair, it may provide greater ethical proof in favor of the regulation of platform's use of Dark Patterns. The next chapter will argue for the case that Dark Patterns are unethical in the sense that they are antithetical to moral values which guide people towards living together.

Chapter 3 - Dark Patterns and Ethics

Chapter 3 Introduction

In this thesis so far, some answers to the research questions have been discussed. First, in chapter 1 an answer was offered to the question: how, if at all, can manipulation change an end-user's values? In short, when some manipulative act succeeds in manipulating a target, it bypasses that person's process of deliberation. If this manipulative act becomes a regular occurrence, then it can be considered to form a part of the manipulated person's regular routine or practice. To the extent that values are shaped in practice, this manipulative act has the potential to shape some of the manipulated person's values to be more in line with those of the manipulator. Second, in chapter 2 an answer was offered to the question: how might manipulation be designed in a technical artifact, and what might that mean for any values embedded or implicated within

that artifact? In short, designers can implement certain choice architectures that exploit cognitive biases which thwart an end user's preferences in an attempt to get them to default to what the designer wants. Since these Dark Patterns are systemically integrated in platforms, they may make up a regular occurrence for the users of platforms. As something regularly occurring in a person's life can be considered to be a part of their routine, Dark Patterns have the potential to shape an end user's values to be more in line with those of the platform. In this final chapter, an answer to the final research question shall be proposed: what are the ethical considerations of any potential harmful consequences of a manipulative digital artifact? In short, this chapter will argue that Dark Patterns have the potential to disrupt ethical decision making for those manipulated by them. To the extent that personal decisions influence societal decisions, Dark Patterns have the potential to not only impact personal decision making, but societal decision making as well. This chapter will conclude that these consequences constitute good reasons to attempt to restrict the use of Dark Patterns by platforms and proposes several methods that may do so.

Section 3.1. Dark Patterns can disrupt ethical decision making

This section argues that a concerning consequence of Dark Patterns is their potential to disrupt ethical decision making. Since Dark Patterns can disrupt decision making in general, as discussed in chapter 2, it stands to reason that ethical decision making may be likewise disrupted. However, this section will argue that there may be some deeper reasons ethical decision making in particular is vulnerable to Dark Patterns. In order to support this claim, this section will make an explicit connection between values and ethics. Singer (2024) describes ethics as a philosophical discipline which studies the concepts of rightness, goodness, and morality. Since values, as defined in chapter 1, regard goodness there is an implicit connection between values and ethics with some exceptions that will be discussed below. It is argued that ethical or moral values are those that help people get along in the collective pursuit of goodness (Gouinlock, 1978). Such collaboration may require mutual understanding which can sometimes be determined through automatic judgment, but may often require reflection and deliberation like, for example, in the case of misunderstanding. To the extent that Dark Patterns encourage fast, automatic, System 1 type thinking and discourage slow, deliberate, System 2 type thinking, they may keep people from some crucial cognitive tools necessary for collaboration in the collective pursuit of goodness. Dark Patterns, therefore, have the worrying potential to disrupt the ability to make ethical decisions. The following two sections elaborate on how this consequence may play out in society.

A moral value can be considered to be a certain kind of value, one that helps determine what is good for a person to do and since what one does in a society is often social, “there is a social aspect to values” (Hickman et al., 2011, p. 176). A certain good valued by an individual is likely to “play out in an interaction with others” (Wieczorek, 2023, p. 361). So, while a certain concept of goodness “might belong to an individual” there is also a collective concept of goodness negotiated as individuals associate with one another (Wieczorek, 2023, p. 361). However, not all values are necessarily ethical which can more specifically be understood to be concerned with what is morally right or good (Singer, 2024). A value that an individual finds good and worthy of pursuit does not necessarily make that value morally right to pursue. Power, for instance, may be a value whose pursuit may result in activities that infringe on others in society. As a mere value, having power may have an allure both instrumentally and for its own sake, but when considering morality power has the potential to disregard others in a given social setting. When power is pursued without regard for others then the value holder might sometimes utilize means regarded as ‘ruthless’ or ‘dictatorial’. These terms are not inconsistent with power as a value but carry a selfish or uncaring connotation and therefore an immoral or unethical significance (Grenberg, 2007). What this example means to show is that good morals are narrower in scope than values. The specific values that help determine what a person will do in order to live in a collective society may be considered morally good and therefore ethical (Gouinlock, 1978; Hickman et al., 2011).

Resolving ethical problems that come up in society likely requires the cognitive faculties of reflection and deliberation as well as relational faculties like empathy (Frankish, 2010; Betzler, 2019; Wieczorek, 2023). Due to its social role, ethics may be considered to be a collaborative effort. While it is possible that some collaborations might be fulfilled by relying only on automatic judgment, more likely there will be at least some issue that arises in the collaborative process that automatic judgment fails to resolve. To the extent that a collaboration is ethically minded and therefore guided by moral values, any conflict that arises in the collaborative process may be a moral problem. If this moral problem cannot be resolved through automatic and intuitive problem-solving then more careful considerations would need to be made if the parties involved wish to resolve the problem. As chapter 1 discussed, instances of failure call for the suggestion of alternative solutions which may require the faculties of reflection and deliberation and the same is true for ethical issues (Dewey and Tufts, 1936). Collaborative reflection and deliberation require a conscious effort to understand why one may think or feel differently from another (Frankish,

2010). Such an attempt at relational understanding may be thought of as a form of empathy (Betzler, 2019).

Since Dark Patterns interfere with reflection and deliberation, as discussed in chapter 2, they may also interfere with the ability to do relational work such as empathizing which, in turn, may make ethical problem-solving more difficult. Living together in a society is marked by the ability to recognize what is meaningful to one another, the values that are important to each other, and how each person has freely made their own lives (Susser et al., 2019). Such empathetic recognition may be slow and effortful, particularly if the other perspective one attempts to recognize is unfamiliar or somehow foreign (Betzler, 2019). As such, this ability to recognize each other in this way may be thought of as sitting firmly within System 2 cognition.

Conversely, some cognitive psychologists have depicted bias and prejudice as sitting within the implicit and automatic judgments of System 1 (Frankish, 2010). This account of bias thinks of it as something that hinders impartial judgment and as such it is counter to the considered judgment of reflection and deliberation. Likewise, bias can be considered opposed to the account of empathy presented above (Dewey and Tufts, 1936; Tschaepe, 2018). As a rigorous attempt to understand someone or something different to oneself, one can use empathy to assist in their collaborative efforts. However, to the extent that someone remains stuck within their bias, it may make thoughtful judgments like empathy more difficult to practice. When bias interferes in this way, it might hinder someone's collaborative efforts. As long as such a bias or prejudice remains unchallenged, a collaborative effort that has run into issues due to such judgments is unlikely to succeed in resolving the issue. To the extent that Dark Patterns keep users within System 1 and out of System 2, it may be reasonably stated that they make acting ethically, i.e., making decisions that improve life together, more difficult in the areas they pervade.

When Dark Patterns subvert individual decision making, and multiple individuals within a society are impacted by that, then that may make collaboration all the more difficult. The issue with Dark Patterns' exploitation of System 1 and subversion of System 2 is not that the System 1 processes are necessarily wrong or bad. The issue is that when System 1 processes go wrong or produce bad results then System 2 may be necessary to correct these. If individual deliberation requires that a single person have access to their System 2 cognitive faculties then ethical decision making likely requires a collaboration between decision-makers with access to System 2. Ethical decisions may require collaborative problem-solving to find moral solutions and considered efforts

such as empathy are important ingredients in that process. Dark Patterns are problematic to the extent that they hinder people's ability to collaborate and considerately relate to one another. The next two sections detail some deeper consequences of this issue. First, Dark Patterns may result in the uncritical acceptance of amoral/immoral values which may make certain ethical practices more difficult. Second, Dark Patterns may inhibit moral growth by hindering an important way people adapt and learn.

Section 3.2. Dark Patterns can promote unethical values

This section argues that the potential of Dark Patterns to undermine ethical decision making may encourage the acceptance of unethical values which may happen in two different ways. First, Dark Patterns have the potential to promote the values of the platforms that utilize them which are not always the values of their users as explained in chapter 2. This section argues that a platform's values are often those which align with shareholder values which may be unethical if they disrupt moral values for the sake of personal gain. Second, Dark Patterns may make it difficult to challenge unethical values associated with the automatic behavior of System 1 cognition. Prejudice, stereotyping, and bias are typically quick and often automatic judgments associated with System 1 cognition (Frankish, 2010). This section argues that these heuristics promote values like supremacy and epistemic authoritativeness which are unethical in function (Tschaepe, 2018). These consequences are potentially harmful effects of Dark Patterns that deserve critical attention.

Dark Patterns may get the users of a given platform to uncritically or implicitly accept the values of that platform as well as their definitions of said values which may, in turn, shape the practice of societal values. As chapter 1 argued, values are vulnerable to influence and manipulation. Furthermore, the choices one makes can play a crucial role in the practices of their values. So, if manipulation is successful in changing a person's choices to be something other than they would have made then it will have, in some way, changed the context of that decision-maker. Such a context is meant to reflect the interests and values of the manipulator (Sunstein, 2016). This context is not only relevant to the individual practices of values, but may also be relevant to any other person who comes into contact with such an individual. The meaning of values can take shape in everyday interactions, as argued in chapter 1 (Wieczorek, 2023). If a person were to be manipulated into acting in accordance with a manipulator's values, then anyone else who interacts with that person may be inadvertently experiencing these values as defined by the manipulator.

In other words, any influence Dark Patterns have on an individual's values might reverberate in the everyday interactions these individuals have with other members of society.

This effect may result in the societal engagement of values as defined by platforms and their shareholders which may be unethical. Manipulation of a digital interface via Dark Patterns "may result in the conversion of users towards shareholder-defined outcomes" including values (Gray et al., 2018, p. 10). If the values promoted by a platform oppose the cooperative pursuit of goodness, for example, then platforms may be encouraging their users to behave unethically. In a study by Adams et al. (2011), the authors make the case that corporations are more likely to side with their shareholders than other stakeholders in any shareholder-stakeholder conflict. The values favored in these conflicts tended to be individualistic ones like power and achievement over cooperative values such as equality and respect (Adams et al., 2011). Power and achievement, in such contexts, are amoral in the sense that their pursuit does not necessarily concern right and wrong. However, if they are pursued to the detriment of moral values as Adams et al. (2011) suggest, then that pursuit may be unethical. If the digital interfaces that subsume user values in favor of shareholder values are distributed throughout society, as chapter 2 argues they are, then the impact of manipulation on personal values may compound into a societal issue. To the extent that platforms do promote individualistic values over moral values, they may be promoting unethical practices in their user base and society at large.

Dark Patterns may also promote implicit values associated with System 1 cognitive processes which may have a tendency to be more tribalistic. Tribalism is here understood to be an attitude of exclusivity that favors one's own group to the detriment of other groups. As this thesis has argued, Dark Patterns may encourage people to favor automatic judgments. These automatic judgments may be inappropriate in the sense that they lack coherent evidence, for instance, requiring a slower deliberate process to correct. Failure to do so may result in these automatic judgements becoming persistently problematic. Certain automatic judgments may favor specific attitudes aligned with certain values. In cognitive psychology, stereotyping and prejudice are sometimes associated with the automatic processes of System 1 (Frankish, 2010). Cognitive biases like prejudice affirm values such as supremacy or superiority which promotes an 'us vs. them' attitude antithetical to moral values aimed at right action in social settings (Ladegaard & Cheng, 2014). As a result, when Dark Patterns exploit cognitive biases they may also have a tendency to promote implicit values of which some may be unethical. Any immoral values

automatically adopted may be all the more difficult to correct so long as Dark Patterns continually succeed in keeping their users thinking or behaving automatically.

This effect may result in platforms being a space that is more conducive to exclusive or tribalistic values rather than cooperative ones. A platform's algorithm often operates automatically, as explained in chapter 2. Furthermore, this automatic operation is often considered 'black-boxed' or inscrutable in terms of function even to the algorithm's engineers (Christin, 2020). Therefore, if the focus is on exploiting cognitive biases to get users to accept the decisions of a platform, an algorithm may not be able to distinguish between the kinds of cognitive biases it capitalizes on. Exploiting a user's sunk cost fallacy and exploiting a user's prejudice may be all the same to an algorithmically programmed Dark Pattern. As a result, platforms may be unknowingly encouraging tribalistic values. There is some empirical evidence that supports this claim. Manacorda et al. (2022) suggests that the rise of social media platforms also saw rises in tribalistic attitudes both online and off. The mechanisms of Dark Patterns explained in this thesis help to explain why this may be the case.

If Dark Patterns exploit cognitive biases to manipulate users into consistently acting automatically as a platform desires, then these platforms might reasonably be promoting individualistic and tribalistic tendencies. Furthermore, if this practice of manipulation adjusts not only the behavior of users but their values as well and those users in turn practice those values in other contexts, then Dark Patterns constitute a societal issue, both online and off. However, there may exist a yet deeper issue where users of platforms may not only be uncritically adopting unethical practices, but may have a more difficult time reflecting and deliberating ethically when regularly exposed to Dark Patterns.

Section 3.3. Dark Patterns can hinder ethical growth

This section argues for one final consequence of the potential of Dark Patterns to undermine ethical decision making. Insofar as Dark Patterns keep users within automatic decision making they may inhibit self-guided personal growth as well as collectively decided societal development. In keeping users within System 1 cognition and out of System 2, this section argues that Dark Patterns restrict or diminish cognitive access to a major faculty used for learning and adaptation. This may further hinder the ability to autonomously make one's own life as deliberated with personal preferences and values. People likely need both System 1 and System 2 cognitive

processes to learn, adapt, and ultimately to grow in their ethical capacities (Frankish, 2010). Furthermore, this section argues that society may need their participants to have sufficient access to both System 1 and System 2 cognitive processes if they are to collaborate and pursue some notion of a collective good and moral value. In particular, resolving the interpersonal conflicts that inevitably ensue in attempts to develop and maintain a society require slow, effortful, and deliberate decision making for all relevant stakeholders. This consequence is a potentially harmful effect of Dark Patterns and deserves critical attention.

Dark Patterns may have the potential to be a hindrance to personal growth. The ability to solve problems is considered necessary for learning or growth in the pragmatic tradition. To resolve a problem encountered and consequently develop one's ability to problem-solve often requires the examination and selection of potential alternative solutions (Hickman et al., 2011). Accordingly, growth and learning often arise through repeated attempts to reflectively adapt to problematic situations that arise due to change (Wieczorek, 2023). This growth, as mentioned in chapter 1, often has a goal in mind which may be directed by personal values. The ability to recognize solutions, other than ones formed through automatic judgment, may require a person to recognize that their preferred solution might not work or that their values potentially need adjusting (Tschaepe, 2018). This form of analytical reflection is of the kind associated with System 2 cognition (Frankish, 2010). While learning and adaptation is possible under both System 1 and System 2, the former, though quicker to access, is often considered to foster slower learning. System 2, on the other hand, is considered to be much more efficient for learning. To the extent that Dark Patterns keep users within System 1 and out of System 2 they encroach upon a substantial portion of a person's ability to develop their reflective and deliberative capacities. As a result, Dark Patterns may not only be regularly overriding someone's personally deliberated values, but they may also be hindering their ability to learn how to reflect on and deliberate their values.

This adverse effect of Dark Patterns on personal growth may extend to the societal ability to develop ethically as they hinder two relevant factors of societal growth. First, societal growth depends on, in part, the ability of its members to learn how to reflect on and potentially adjust their values. Second, societal growth depends on, in part, the ability to resolve interpersonal conflicts. These may not be the only ingredients for social development, but they are important and relevant concerning the effects of Dark Patterns. The following paragraphs will explain how these two challenges are relevant to ethical decision-making.

Dark Patterns have the potential to interfere with social structure and change. The basic reasoning is that a society is made up of individuals who organize their social structure through their interactions (Form & Wilterdink, 2018). When a social structure changes, it can lead to a more permanent social change. Furthermore, an “individual’s sense of self [often] reflects their socialized experiences, relationships, and other internalized aspects of social structure” (Hitlin & Long, 2009, p. 138). When Dark Patterns regularly influence an individual's decisions it has the potential to also change their sense of self as developed in the context of manipulation. What someone likes, what they value, and what they find important are all concepts which may contribute to their sense of self which may change insofar as Dark Patterns change their likes, values, etc. If this change has an impact on one’s relationships and social experiences, then that personal change may have an impact on their social structure. For example, if Dark Patterns succeed in undermining a user’s value of privacy by getting them to give up their data, this change may alter this individual's social structure through the ways they interact in society as a result of that change. If this individual’s perspective on privacy has changed, they may reiterate that perspective in their societal interactions which may influence their social structure. When Dark Patterns interfere with personal growth, they may be inadvertently directing societal change.

Dark Patterns also have the potential to make resolving moral problems more difficult. Resolving conflict often requires conscious reflection and deliberation in order to understand the situation, examine any relevant values, and consider possible solutions (Gouinlock, 1978). Such judgments require one to be conscious about not only the relevant facts at hand, but their own judgments as well (Tschaeppe, 2018). Dark Patterns may not only inhibit understanding why others may think differently from one another, as argued in section 1 of this chapter, they may also inhibit someone understanding what they themselves think and why. When Dark Patterns keep people in System 1 and out of System 2, they may also be inhibiting their ability to scrutinize their own reasons and values. If someone is less capable of scrutinizing a change to their personal values, then they may carry this limitation with them in any social interactions they may have. Should that change become problematic and cause conflict then the lack of scrutiny may make finding a solution all the more difficult. As Dark Patterns are a systemic problem, as argued in chapter 2, there may be many different platform users whose scrutiny may be impaired, making interpersonal conflict resolution more difficult on a societal level. As the first section of this chapter discussed, a conflict that arises in a collaborative process of determining right action may be considered a moral problem. Resolving moral problems is important for the ethical development of any society.

Therefore, a society that is hindered in its ability to resolve conflicts is hindered in its ability to develop ethically (Gouinlock, 1978; Wieczorek, 2023).

In the conclusion of their article, Luguri and Strahilevitz (2021) propose that, although “manipulation in the marketplace is a longstanding problem”, Dark Patterns have exacerbated the problem greatly (p. 104). While the focus of researchers into Dark Patterns has predominantly been on the legal and immediate impacts of Dark Patterns, this thesis has attempted to provide a philosophical grounding for why Dark Patterns are problematic. The tendency of Dark Patterns to exploit cognitive biases to manipulate a user into automatically accepting the preferred decision of a platform is what makes them uniquely problematic. Not only are they experienced negatively in the ways they violate a person’s autonomy and self-chosen values, but Dark Patterns may also interfere with a user’s moral growth and entrench ethically problematic behavior such as prejudice and bias. These personal issues may in turn result in deep-seated societal issues as well. The following section proposes potential resolutions that may help limit the uses of Dark Patterns in UX design and alleviate some of their harmful effects.

Section 3.4. Some ethical suggestions

The problems of Dark Patterns argued above may be considered sufficient reasons to call for ethical solutions and, accordingly, this section will propose several. I propose three potential and partial solutions that seek to minimize the ethical harms of Dark Patterns as well as promote appropriate legal repercussions and reasonable alternatives. The three suggestions proposed are as follows: (1) resistance (both personal and societal including legal measures), (2) awareness and transparency, and (3) promoting design solutions for platforms that are Dark Pattern free. This section will provide arguments and evidence for the effectiveness of these measures as well as potential advantages and drawbacks of each.

3.4.1. Resistance

Depending on how much they interfere with a person’s life, an ideal may be to resist Dark Patterns wherever warranted and whenever possible. This might be accomplished through several means including minimizing time spent on platforms, installing software made to mitigate certain Dark Patterns, and forming new policies to restrict their usage in design. Firstly, the less time a user spends on platforms, the less exposure to Dark Patterns they will have. To that end, there is some wisdom in the colloquial meme ‘go touch grass’ which, in short, means to put away devices and

go out and experience the 'real' world. Doing so may offer some meaningful health benefits including improving judgment and lowering risks for stress-induced health complications (Price, 2019). End users can also limit their exposure to platforms without separating from technology entirely. Disabling notifications and using non-platformed apps and devices (such as e-books, offline Blu-ray Players, and even some video games) may also provide some meaningful distance from Dark Patterns. However, such opportunities may be few and far between for workers with a company device meant to be close at hand at all times during work hours. Given the systemic nature of Dark Patterns, some companies may need to reconsider such policies for the wellbeing of their employees.

Another method of resistance is to install third party software designed to bypass certain Dark Patterns. For instance, Ghostery is a company that has developed a few software applications with an emphasis on privacy and transparency (Ghostery Manifesto, n.d.). Their flagship ad-blocking browser plugin has a feature that allows a user to automatically refuse consent to platforms asking for personal data whenever possible. Such a feature bypasses many different Dark Patterns that might be used to coerce a user into giving consent. While no complete list of anti-dark-pattern software currently exists, the Berlin based Tactical Tech does have a site listing alternative apps aimed at improving individual privacy (Tactical Tech, n.d.). While these apps are not specifically anti-dark-pattern, their pro-privacy mission is complementary to an anti-dark-pattern mission and a good place to start for any individual looking to reduce their number of Dark Patterns encountered. Future research might create a catalog of software solutions meant to resist Dark Patterns.

Policies can be made and enforced to actively resist Dark Patterns. There are some legal policies in place that theoretically restrict the usage of Dark Patterns. The EU has officially prohibited the use of Dark Patterns through the Digital Services Act (DSA). However, as of 2023, there is still a question of how to implement and enforce this prohibition (Akhurst et al., 2023). In the U.S., forced continuity, roach motel, hidden information, disguised ads, aesthetic manipulation, bait-and-switch, and other uses of Dark Patterns (see Table 1 in chapter 2) have all been penalized in cases plaintiffed by the Federal Trade Commission (FTC). These cases, according to Luguri & Strahilevitz (2021) suggest that legally restricting the more harmful Dark Patterns may be substantially achievable under existing laws (within the U.S. for these examples). However, given their still pervasive use on the world's most popular platforms, more should be done (Bongard-Blanchy et al., 2021; Mildner et al., 2023). One suggestion, by Luguri & Strahilevitz (2021), is to

give existing legislation more teeth by empowering agencies like the FTC to perform counter-A/B testing on platforms suspected of using Dark Patterns. These tests, they argue, could be used to “track user responses to ensure that consumers’ choices are broadly consistent with their preferences” (p. 99). Such measures are, however, reactive and so may not do enough to prevent the use of Dark Patterns. This is especially the case for companies who believe they can get away with unethical decisions, who include the price of punitive measures in their cost of doing business, or who avoid compliance by basing their company in a country without such standards.

However, policy makers also have a lot of room to create new and improve existing measures focused on resisting Dark Patterns. These measures include advocacy as proposed by Gray et al. (2021), enforcing user-friendly options to reject web tracking (Borberg et al., 2022), decentralizing online social spaces (Westin and Chiasson, 2021), and creating “economic incentives and regulatory interventions” that work to eliminate Dark Patterns from online platforms (Bongard-Blanchy et al., 2021, p. 773). These authors argue that such forms of resistance will likely have a significant impact on reducing the usage and harm of Dark Patterns. These new policies will need to be carefully considered if they are to succeed in facilitating compliance as platforms may determine that Dark Patterns are still worth using (Armour et al., 2020).

3.4.2. Awareness and Transparency

One of the issues with Dark Patterns is that they are sometimes subtle or deceptive. Many end users are unaware of the existence of Dark Patterns which makes resisting them rather difficult (Pot, 2023). Strategies which focus on cultivating the popular awareness of Dark Patterns may help to reduce their influence. Several studies provide evidence that generating awareness may “strengthen people’s resistance to [Dark Patterns]” (Bongard-Blanchy et al., 2021, p. 774; see also Westin and Chaisson, 2021 and Luguri & Strahilevitz, 2021). The reasoning for this is simply that to avoid danger one may need to be aware of it. A lack of awareness is often a key ingredient for Dark Patterns to function, especially the more deceptive varieties that aim to trick or mislead (Borberg et al., 2022). Raising awareness can give language to the negative ways Dark Patterns impact users as well as provide information on the tools available to contend with Dark Patterns. However, these studies also argue that awareness is often not sufficient to avoid being manipulated by Dark Patterns (Bongard-Blanchy et al., 2021; Westin and Chiasson, 2021; Mildner et al., 2023). As covered in chapter 2, users may remain manipulated despite being aware of the attempts of Dark Patterns to manipulate them. So, though awareness may be a necessary ingredient, it may not be sufficient to counter the negative effects of Dark Patterns.

Requiring transparency may also improve this systemic problem. Transparency here means to make plain or to disclose. The covertness of some Dark Patterns can make them difficult to detect and the normalization of Dark Patterns may increase as users get used to them. Legally or officially recognizing Dark Patterns and requiring platforms to disclose their use may undo some of this obfuscation. There is precedence for similar legislation having a positive impact. The success of the General Data Protection Regulation (GDPR) and the approval of the recent EU Artificial Intelligence (AI) Act demonstrate that the EU can have substantial sway over enforcing ethical mandates for digital technologies (Zaeem & Barber, 2021; Artificial Intelligence (AI) Act, 2024). The EU may be able to follow in the footsteps of these like-minded legal frameworks to create a new policy that requires the disclosure of the use of legally recognized Dark Patterns.

Furthermore, such legislation may help to enforce existing policies like the GDPR which Dark Patterns have been shown to undermine (Nouwens et al., 2020, see Appendix A). However, this does have the potential for a downside similar to what happened after the advent of the GDPR. Since its implementation, end users have had to deal with a substantial increase in the number of popups concerning privacy and consent which they must deal with before they can use a website or app. The less a user consents, the more popups they may need to deal with causing frustration and confusion (Burgess, 2018). Regulating Dark Patterns by requiring transparency may follow a similar misfortune. Legislators will need to think about how to enforce such a law in ways that do not tax their constituents too greatly so as to defeat the purpose of such a law.

3.4.3. Alternative design patterns

Finally, companies and public agencies may be able to encourage decreasing the use of Dark Patterns through the promotion of alternative design patterns which are opposed to the design philosophy of Dark Patterns. I shall discuss two such methods known as Bright Patterns and useful friction. Bright Patterns attempt to do exactly the opposite as Dark Patterns but use the same kind of design strategy. Specifically, Bright Patterns are still a form of digital manipulation, but promote options which are meant to favor the user as opposed to the platform (Bongard-Blanchy et al., 2021). For example, a consent form designed using Bright Patterns would make it easier for a user to deny consent and more difficult for them to opt in. While Bright Patterns may prioritize certain consumer protections, many of the same harms as identified earlier in this thesis remain. As they are still manipulative, Bright Patterns may still promote “unreflective default [behavior] and users’ perception of a lack of control” (Graßl et al., 2021, p. 25).

Another design pattern opposed to Dark Patterns is known as useful friction which focuses on improving user choice by adding friction to the end-user experience (Westin and Chiasson, 2021, p. 10). Popularized by Tristan Harris, the co-founder of the Center for Humane Technology, useful friction is an idea that reacts to the sense that platforms have become too easy to navigate, i.e., too frictionless. However, not every choice is equally easy. Platforms “*want to make the choices they want you to make easier, and the choices they don’t want you to make harder*” and, this thesis has argued, use Dark Patterns to achieve this (Harris, 2016, original emphasis). In other words, one way to think about the manipulateness of Dark Patterns is that they can add an overabundance of friction to choices platforms want their users to reject. Furthermore, they might add too little friction to the choices platforms want their users to accept.

Westin and Chiasson (2021) argue that adding an appropriate amount of friction to design patterns can make an end user's actions clearer as they must slow down to consider their options more carefully. For example, if an online store wishes to sign a user up for a free trial they could add equal measures of friction to both the option to accept and decline. Such a system designed with the dark pattern ‘forced continuity’ (see Table 1 in chapter 2) would attempt to get the user to accept the free trial implicitly and auto-renew at full price once the trial expired. That same system designed with useful friction may require the user to manually select if they would like to sign up for such a trial or not and, if so, if they would like to auto-renew. The main idea is that each option presented to a user should be relatively similar in terms of the effort required to select it. Such a design solution is in line with encouraging System 2 thinking, which is again slower, considered, and conscious. Thus, useful friction also falls in line with the pragmatic inclinations of this thesis in that friction may help one slow down and consider how their values may or may not apply to the situation at hand to help resolve it..

The suggestions in this section are meant to counteract some of the harms that Dark Patterns can cause by offering some ways to diminish exposure to Dark Patterns, increasing the capacity to identify instances them, and proposing some alternatives to Dark Patterns for future UX design. No single suggestion is a guaranteed fix to the problem at hand, and any wholistic approach to resolving this societal problem will likely need some combination of methods to succeed. Furthermore, this thesis is limited to a predominantly theoretical argument, no original empirical research was performed which is something future research may wish to undertake. However, if Dark Patterns are a problem worth resolving, then these are some ideas to get started.

Conclusion

Dark Patterns are software solutions which have incorporated regular and systemic use of manipulation to habitually steer the practices of their users towards the aims of a platform (Choudary, 2015; Susser et al., 2019). When Dark Patterns bypass or undermine considered and conscious thinking, they have the potential to diminish a person's ability to reflect on, develop, or adapt their values. To the extent that Dark Patterns are enactments of systemic manipulation, they have the worrying potential to prevent individuals as well as society from making decisions based on their values.

Values are concepts that concern goodness, as this thesis has argued, and may require conscious, reflective, and deliberative thinking if a person is to intentionally direct their actions towards that goodness. In practice, values might help guide such action in the pro-attitudes they can inspire, i.e., attitudes that fit or align with what someone wants or values, like fairness possibly inspiring an impartial attitude for example (Svavarsdóttir, 2014). However, values are not necessarily static concepts and the pro-attitudes they inspire may change in accordance with new and unique circumstances that can occur in practice (Dewey, 1922b). As a result, a person's values may be open to change both in meaning and relative importance. On the one hand, this openness is a good thing in the sense that what matters to people can remain important as they figure out how to keep or change their values with respect to a new situation. In this way, the meaning of values is continually reflected on and developed in practice (Boenink & Kudina, 2020). On the other hand, because of the ways values take shape in practice, it is possible that values are also sensitive to influences within practice, including the influential power of manipulation (Wieczorek, 2023).

Dark Patterns are manipulative, as this thesis has argued, and have a tendency to exploit implicit biases (Luguri and Strahilevitz, 2021). An objective of platforms is what some in the software industry call 'behavior design' which seeks to keep and maintain a user's attention in order to potentially convert that attention into a profitable behavior (Choudary, 2015). The design philosophy of platforms has heavily relied on continuous experimentation to find the most effective ways of getting a user to do as the platform wishes (Quin et al., 2024). This technique has developed ways of turning attention into behaviors which benefit the platform most. These Dark Patterns are effective in manipulating users into accepting what the platform has determined is best for their business rather than what the user may themselves think to be right and good (Luguri

& Strahilevitz, 2021). This kind of acceptance is something these platforms try to automate, which makes Dark Patterns a kind of technology that can be difficult to keep up with. While attempts at mass manipulation are nothing new (Bakeless, 1931), as with the use of deception in advertising discussed in chapter 2 for example, the scale and speed of manipulation that comes with Dark Patterns is new.

This raises cause for concern. We may be entering an era when people who spend their time regularly on platforms have less cognitive space to think about and practice their values, either personally or collectively. One's values, as argued in chapter 1, can be formed and maintained through reflection and deliberation. Dark Patterns, as argued in chapter 2, undermine reflection and deliberation when they keep users within automatic thinking and away from conscious thinking. Reflection and deliberation, as argued in chapter 3, can also be important for ethical decision making. Meanwhile, the amount of time users spend on social media platforms continues to rise, especially since the COVID-19 pandemic (Trott et al., 2022). End users who spend a significant amount of time on platforms are regularly exposed to the Dark Patterns those platforms use, as this thesis has argued. With that much time exposed to rapidly-evolving and effective methods of manipulation, these users may be living more automatically and less consciously than they otherwise might. If so, then it may be more difficult for these users to reflect on or develop their values in uncertain circumstances, personally or socially.

If society is to be able to evaluate the impact technology has on it, then it likely needs its members to have the ability to evaluate their values. When people are hindered in their access to reflecting and deliberating on the things that matter to them, then it may become difficult for them to communicate this to others. If enough people have trouble communicating what matters to them, then it may become difficult for their society to collectively decide what matters to it as a whole. If a substantial number of people are under the influence of mass manipulation, then it is possible that their society might develop to be more in favor of what matters to the manipulators rather than what matters to that society. Considering how widely and heavily used digital platforms are, this thesis proposes that this may be precisely what platforms are doing through their extensive use of Dark Patterns (Langley and Leyshon, 2017; Giraldo-Luque et al., 2020; Westin and Chiasson, 2021). In other words, there is a serious risk that platforms are manipulating society to conform to the platforms' aims and values through the use of software design methods that are highly effective, rapidly modifiable, and black-boxed or inscrutable (Luguri & Strahilevitz, 2021; Christin, 2020).

Some solutions are presented in this thesis in chapter 3, such as promoting methods of resistance or the creation of new legislation. However, I propose more work needs to be done in three domains specifically. First, though the term Dark Patterns is entering into the mainstream, it is still a niche of which many remain unaware (Pot, 2023). As discussed in chapter 3, while awareness is not sufficient for reducing the effectiveness of Dark Patterns, it may be necessary. Given the ethical consequences of Dark Patterns, more should be done by academics and policy makers to disseminate this knowledge (Bongard-Blanchy et al., 2021; Westin and Chaisson, 2021; Luguri & Strahilevitz, 2021).

Second, more policy research is needed that focuses on preventing Dark Patterns from undermining policy goals. Dark Patterns might allow platforms to adhere to the letter of the law while subtly ignoring the spirit of the law, as they may be doing in response to the GDPR as discussed in chapter 3 (Burgess, 2018). There is a gap in research on best practices for implementing policies that are less vulnerable to Dark Patterns. Filling that gap would be beneficial to strengthening current and future policies regarding the tech industry.

Third and finally, 'useful friction', i.e., adding equal and appropriate digital constraints to all of the relevant choices in a software interface, is a little studied concept that has some exciting synergy with the research presented in this thesis. Intentionally adding friction to software solutions may have the benefit of making the options presented clear, encouraging users to slow down and think more carefully about their choices (Westin and Chiasson, 2021). Such slow and considered thinking can help a decision-maker to make choices that matter to them, rather than automatic or habitual choices that matter to some platform. Of course, useful friction may have some drawbacks. For instance, it may not be desirable for users who already practice minimal use of social media and for whom the added friction may cause them to stop using social media altogether. However, to the extent that a user spends a significant amount of time on social media, it may benefit them to take their time and try to use that time wisely (Graciyal & Viswam, 2021).

Given the above arguments, I propose a possible answer to the main research question asked at the outset of this thesis. Why, how, and to what extent are Dark Patterns utilized by an online social media platform to manipulate end-user behavior and values in the service of that platform's commercial interest and what may be the harm of such manipulation? A platform's values are not explicitly used to manipulate, at least not insofar as the research put forward in this thesis

demonstrates. Rather they can be considered to be favored when the platform undermines the values of their users through manipulative tactics. The aim is to capture and keep a user's attention so they might turn that attention into behaviors which commercially benefit the platform. This process may simultaneously promote the platform's values while also undermining their users' values. This likely happens to the extent that Dark Patterns are systemically used which, as argued, is significantly so. For those who wish to use their cognitive faculties of reflection and deliberation to help them live in accordance with their values, Dark Patterns may present a substantial obstacle. This obstacle is important to understand for any person or society that wishes to consciously progress in the ways that matter to them.

References

- Adams, R. B., Licht, A. N., & Sagiv, L. (2011). Shareholders and stakeholders: How do directors decide? *Strategic Management Journal*, 32(12), 1331–1355. <https://doi.org/10.1002/smj.940>
- Akhurst, T., Zurdo, L., Rapparini, R., & Markhof, C. M. (2023, April). *How should the European Union regulate dark patterns?*
- Armour, J., Gordon, J., & Min, G. (2020). Taking Compliance Seriously. *Yale Journal on Regulation*, 37.
- Artificial Intelligence (AI) Act: Council Gives Final Green Light to the First Worldwide Rules on AI. (2024, May 21) *Council, European Union - Consilium*, Council of the EU, www.consilium.europa.eu/en/press/press-releases/2024/05/21/artificial-intelligence-ai-act-council-gives-final-green-light-to-the-first-worldwide-rules-on-ai
- Bakeless, J. (1931, March). Machine-made minds: the psychological effects of modern technology. *MIT Technology Review*, 33(6), 275–300.
- Baldrige, R. (2022, April 18). *Understanding Venture Capital*. Forbes Advisor. <https://www.forbes.com/advisor/investing/venture-capital/>
- Betzler, M. (2019). The Relational Value of Empathy. *International Journal of Philosophical Studies*, 27(2), 136–161. <https://doi.org/10.1080/09672559.2019.1598081>
- Boenink, M., & Kudina, O. (2020). Values in responsible research and innovation: From entities to practices. *Journal of Responsible Innovation*, 7(3), 450–470. <https://doi.org/10.1080/23299460.2020.1806451>
- Bongard-Blanchy, K., Rossi, A., Rivas, S., Doublet, S., Koenig, V., & Lenzi, 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. <https://doi.org/10.1145/3461778.3462086>

- Borberg, I. M., Hougaard, R., Rafnsson, W., & Kulyk, O. (2022). "So I Sold My Soul": Effects of Dark Patterns in Cookie Notices on End-User Behavior and Perceptions. *Proceedings 2022 Symposium on Usable Security*. Symposium on Usable Security, San Diego, CA.
<https://doi.org/10.14722/usec.2022.23026>
- Borchers, N. S., & Enke, N. (2022). "I've never seen a client say: 'Tell the influencer not to label this as sponsored'": An exploration into influencer industry ethics. *Public Relations Review*, 48(5), 102235. <https://doi.org/10.1016/j.pubrev.2022.102235>
- Brignull, H. (2010). Deceptive Design. *Www.deceptive.design*. <https://www.deceptive.design/>
- Burgess, M. (2018, August 29). The tyranny of GDPR popups and the websites failing to adapt. *Wired*. <https://www.wired.com/story/gdpr-cookies-epivacy-regulation-popups/>
- Choudary, Sangeet Paul, Parker, G., & Marshall Van Alostne. (2015). *Platform scale: how an emerging business model helps startups build large empires with minimum investment*. Platform Thinking Labs.
- Christin, A. (2020). The ethnographer and the algorithm: Beyond the black box. *Theory and Society*, 49(5–6), 897–918. <https://doi.org/10.1007/s11186-020-09411-3>
- Cohen, S. (2023). Are All Deceptions Manipulative or All Manipulations Deceptive? *Journal of Ethics and Social Philosophy*, 25(2). <https://doi.org/10.26556/jesp.v25i2.1998>
- De Monticelli, R. (2019). The Paradox of Axiology. A Phenomenological Approach to Value Theory. *Phenomenology and Mind*, 116-128 Pages. https://doi.org/10.13128/PHE_MI-24976
- Dewey, J. (1922). *Human nature and conduct: An introduction to social psychology*. Henry Holt and Company. <https://doi.org/10.1037/14663-000>
- Dewey, J. (1922). Valuation and Experimental Knowledge. *The Philosophical Review*, 31(4), 325. <https://doi.org/10.2307/2179099>
- Dewey, J. (1925). The Meaning of Value. *The Journal of Philosophy*, 22(5), 126. <https://doi.org/10.2307/2014306>

- Dewey, J., & Tufts, J. (1932). *Ethics* (Revised Edition). Holt, Rinehart and Winston, Inc. (Original work published 1908)
- Division of Financial Practices. (2022, September 15). *Bringing Dark Patterns to Light | Staff Report*. Federal Trade Commission; FTC BUREAU OF CONSUMER PROTECTION.
<https://www.ftc.gov/reports/bringing-dark-patterns-light>
- Ersche, K. D., Ward, L. H. E., Lim, T.-V., Lumsden, R. J., Sawiak, S. J., Robbins, T. W., & Stochl, J. (2019). Impulsivity and compulsivity are differentially associated with automaticity and routine on the Creature of Habit Scale. *Personality and Individual Differences, 150*, 109493.
<https://doi.org/10.1016/j.paid.2019.07.003>
- Fields, D., Bissell, D., & Macrorie, R. (2020). Platform methods: Studying platform urbanism outside the black box. *Urban Geography, 41*(3), 462–468.
<https://doi.org/10.1080/02723638.2020.1730642>
- Form, W., & Wilterdink, N. (2018). social structure | Definition, Examples, Characteristics, & Facts. In Encyclopædia Britannica. <https://www.britannica.com/topic/social-structure>
- Frömer, R., & Shenhav, A. (2022). Filling the gaps: Cognitive control as a critical lens for understanding mechanisms of value-based decision-making. *Neuroscience & Biobehavioral Reviews, 134*, 104483. <https://doi.org/10.1016/j.neubiorev.2021.12.006>
- Frankish, K. (2010). Dual-Process and Dual-System Theories of Reasoning. *Philosophy Compass, 5*(10), 914–926. <https://doi.org/10.1111/j.1747-9991.2010.00330.x>
- Ghostery Manifesto. (n.d.). Ghostery. Retrieved June 3, 2024, from
<https://www.ghostery.com/ghostery-manifesto>
- Giraldo-Luque, S., Aldana Afanador, P. N., & Fernández-Rovira, C. (2020). The Struggle for Human Attention: Between the Abuse of Social Media and Digital Wellbeing. *Healthcare, 8*(4), 497.
<https://doi.org/10.3390/healthcare8040497>
- Gouinlock, J. (1978). Dewey's Theory of Moral Deliberation. *Ethics, 88*(3), 218–228.
<https://doi.org/10.1086/292073>

- Graßl, P., Schraffenberger, H., Borgesius, F. Z., & Buijzen, M. (2021). *DARK AND BRIGHT PATTERNS IN COOKIE CONSENT REQUESTS*. 3(1).
- Graciyal, D. G., & Viswam, D. (2021). Social Media and Emotional Well-being: Pursuit of Happiness or Pleasure. *Asia Pacific Media Educator*, 31(1), 99–115.
<https://doi.org/10.1177/1326365X211003737>
- Gray, C. M., & Boling, E. (2016). Inscribing ethics and values in designs for learning: A problematic. *Educational Technology Research and Development*, 64(5), 969–1001.
<https://doi.org/10.1007/s11423-016-9478-x>
- Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018, April). The dark (patterns) side of UX design. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1-14).
- 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.
<https://doi.org/10.1145/3479516>
- Grenberg, J. (2007). Dependent and Corrupt Rational Agency. *Kant Studien*, 98(1).
<https://doi.org/10.1515/KANT.2007.004>
- Harris, T. (2016, May 18). How Technology is Hijacking Your Mind—from a Former Insider. Medium; Thrive Global. <https://medium.com/thrive-global/how-technology-hijacks-peoples-minds-from-a-magician-and-google-s-design-ethicist-56d62ef5edf3>
- Hickman, L. A., Flamm, M. C., Skowroński, K. P., & Rea, J. A. (2011). *The Continuing Relevance of John Dewey: Reflections on Aesthetics, Morality, Science, and Society*. BRILL.
<https://doi.org/10.1163/9789042032330>
- Hirose, I., & Olson, J. (Eds.). (2015). *The Oxford Handbook of Value Theory*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199959303.001.0001>

- Hill, S. (2024, June 27). *How to Spot Fake Reviews on Amazon*. Wired.
<https://www.wired.com/story/how-to-spot-fake-reviews-amazon/>
- Hitlin, S., & Long, C. (2009). Agency as a Sociological Variable: A Preliminary Model of Individuals, Situations, and the Life Course. *Sociology Compass*, 3(1), 137–160.
<https://doi.org/10.1111/j.1751-9020.2008.00189.x>
- Jang, S., & Yu, F. (2008). The triad of credentialing processes in Engineer Quality Regulation: Accreditation, certification, and licensure/registration. *PICMET '08 - 2008 Portland International Conference on Management of Engineering & Technology*, 1955–1962.
<https://doi.org/10.1109/PICMET.2008.4599816>
- Kaag, J. (2013). Fallibility and Insight in Moral Judgment. *Human Studies*, 36(2), 259–275.
<https://doi.org/10.1007/s10746-012-9252-y>
- Klenk, M. (2021). How Do Technological Artefacts Embody Moral Values? *Philosophy & Technology*, 34(3), 525–544. <https://doi.org/10.1007/s13347-020-00401-y>
- König, P. D. (2022). Two tales about the power of algorithms in online environments: On the need for transdisciplinary dialogue in the study of algorithms and digital capitalism. *Media, Culture & Society*, 44(7), 1372–1382. <https://doi.org/10.1177/01634437221111893>
- Korsgaard, C. M. (1983). Two Distinctions in Goodness. *The Philosophical Review*, 92(2), 169.
<https://doi.org/10.2307/2184924>
- Ladegaard, H. J., & Cheng, H. F. (2014). Constructing the cultural 'other': Prejudice and intergroup conflict in university students' discourses about 'the other.' *Language and Intercultural Communication*, 14(2), 156–175. <https://doi.org/10.1080/14708477.2013.849718>
- Langley, P., & Leyshon, A. (2017). Platform capitalism: The intermediation and capitalisation of digital economic circulation. *Finance and Society*, 3(1), 11–31.
<https://doi.org/10.2218/finsoc.v3i1.1936>

- Le Dantec, C. A., Poole, E. S., & Wyche, S. P. (2009). Values as lived experience: Evolving value sensitive design in support of value discovery. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1141–1150. <https://doi.org/10.1145/1518701.1518875>
- Luguri, J., & Strahilevitz, L. J. (2021). Shining a Light on Dark Patterns. *Journal of Legal Analysis*, 13(1), 43–109. <https://doi.org/10.1093/jla/laaa006>
- Manacorda, M., Tabellini, G., & Tesei, A. (2022). Mobile Internet and the Rise of Political Tribalism in Europe. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4235470>
- Mildner, T., Savino, G.-L., Doyle, P. R., Cowan, B. R., & Malaka, R. (2023). About Engaging and Governing Strategies: A Thematic Analysis of Dark Patterns in Social Networking Services. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–15. <https://doi.org/10.1145/3544548.3580695>
- Nickel, P. J., Kudina, O., & Van De Poel, I. (2022). Moral Uncertainty in Technomoral Change: Bridging the Explanatory Gap. *Perspectives on Science*, 30(2), 260–283. https://doi.org/10.1162/posc_a_00414
- Nouwens, M., Liccardi, I., Veale, M., Karger, D., & Kagal, L. (2020). Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3313831.3376321>
- Oddie, G. (2015). Value and Desires. In I. Hirose & J. Olson (Eds.), *The Oxford Handbook of Value Theory* (pp. 60–79). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199959303.001.0001>
- Papadopoulos, P., Kourtellis, N., Rodriguez, P. R., & Laoutaris, N. (2017). If you are not paying for it, you are the product: How much do advertisers pay to reach you? *Proceedings of the 2017 Internet Measurement Conference*, 142–156. <https://doi.org/10.1145/3131365.3131397>

- Pattanaik, P., & Xu, Y. (2015). Freedom and its Value. In I. Hirose & J. Olson (Eds.), *The Oxford Handbook of Value Theory* (pp. 356–380). Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199959303.001.0001>
- Pot, J. (2023, July 2). “Dark patterns” can ruin your online experience. Here’s how to avoid them. Popular Science. <https://www.popsci.com/diy/dark-patterns/>
- Price, C. (2019, April 24). Putting Down Your Phone May Help You Live Longer. The New York Times. <https://www.nytimes.com/2019/04/24/well/mind/putting-down-your-phone-may-help-you-live-longer.html>
- Quin, F., Weyns, D., Galster, M., & Silva, C. C. (2024). A/B testing: A systematic literature review. *Journal of Systems and Software*, 211, 112011. <https://doi.org/10.1016/j.jss.2024.112011>
- Rabinowicz, W., & Rønnow-Rasmussen, T. (2004). The Strike of the Demon: On Fitting Pro-attitudes and Value. *Ethics*, 114(3), 391–423. <https://doi.org/10.1086/381694>
- Sagiv, L., & Roccas, S. (2021). How Do Values Affect Behavior? Let Me Count the Ways. *Personality and Social Psychology Review*, 25(4), 295–316. <https://doi.org/10.1177/10888683211015975>
- Schroeder, M. (2021). *Value Theory* (Stanford Encyclopedia of Philosophy) (E. Zalta, Ed.). Stanford.edu. <https://plato.stanford.edu/entries/value-theory/>
- Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*, 4(2), 205395171773810. <https://doi.org/10.1177/2053951717738104>
- Singer, P. (2024). Ethics | Origins, History, Theories, & Applications. In Encyclopedia Britannica. <https://www.britannica.com/topic/ethics-philosophy>
- Smith, D. G. (2023, January 9). How to Focus Like It’s 1990. The New York Times.
- Sunstein, C. R. (2016). Fifty Shades of Manipulation. *Journal of Marketing Behavior*, 1(3–4), 214–244. <https://doi.org/10.1561/107.00000014>
- Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review*, 8(2). <https://doi.org/10.14763/2019.2.1410>

- Svavarsdóttir, S. & Journal of Philosophy Inc. (2014). Having Value and Being Worth Valuing: *Journal of Philosophy*, 111(2), 84–109. <https://doi.org/10.5840/jphil201411125>
- Tactical Tech. (n.d.). Alternative App Centre. Datadetoxkit.org. Retrieved June 3, 2024, from <https://datadetoxkit.org/en/alternative-app-centre/>
- Thies, F., Huber, A., Bock, C., Benlian, A., & Kraus, S. (2019). Following the Crowd—Does Crowdfunding Affect Venture Capitalists' Selection of Entrepreneurial Ventures? *Journal of Small Business Management*, 57(4), 1378–1398. <https://doi.org/10.1111/jsbm.12447>
- Trott, M., Driscoll, R., Iraldo, E., & Pardhan, S. (2022). Changes and correlates of screen time in adults and children during the COVID-19 pandemic: A systematic review and meta-analysis. *eClinicalMedicine*, 48, 101452. <https://doi.org/10.1016/j.eclinm.2022.101452>
- Tschaepe, M. (2018). Cultural Humility and Dewey's Pattern of Inquiry: Developing Good Attitudes and Overcoming Bad Habits. *Contemporary Pragmatism*, 15(1), 152–164. <https://doi.org/10.1163/18758185-01501007>
- Van De Poel, I. (2020). Values and Design. In D. P. Michelfelder & N. Doorn (Eds.), *The Routledge Handbook of the Philosophy of Engineering* (1st ed., pp. 300–314). Routledge. <https://doi.org/10.4324/9781315276502-27>
- Westin, F., & Chiasson, S. (2021). "It's So Difficult to Sever that Connection": The Role of FoMO in Users' Reluctant Privacy Behaviours. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15. <https://doi.org/10.1145/3411764.3445104>
- Wieczorek, M. (2023). Dewey's Notion of Intelligent Habit as a Basis for Ethical Assessment of Technology. *Contemporary Pragmatism*, 20(4), 356–377. <https://doi.org/10.1163/18758185-bja10074>
- Willis, L. E. (2020). Deception by design. *Harv. JL & Tech.*, 34, 115.
- Zaeem, R. N., & Barber, K. S. (2021). The Effect of the GDPR on Privacy Policies: Recent Progress and Future Promise. *ACM Transactions on Management Information Systems*, 12(1), 1–20. <https://doi.org/10.1145/3389685>

Zimmer, W. (1995). Relationships between design patterns.

Zuboff, S. (2015). Big other: Surveillance Capitalism and the Prospects of an Information Civilization.

Journal of Information Technology, 30(1), 75–89. <https://doi.org/10.1057/jit.2015.5>

Appendices

Appendix A - Screenshots of Dark Patterns

This appendix compiles a number of screenshots found within the literature examined for this thesis. While the thesis itself relied on some descriptions of certain Dark Patterns to support the argument, it can be helpful to see the precise interfaces that are being called problematic and unethical. These examples are not comprehensive, but hopefully they can help give a good idea of what an end user can regularly expect to encounter on platforms. This appendix will provide examples from Gray et al. (2018), Luguri & Strahilevitz (2021), Mildner et al. (2023), and Nouwens et al. (2020) in that order.

Examples from Gray et al. (2018)

The aim of Gray et al's (2018) research was to assemble a comprehensive compilation of what had been termed 'Dark Patterns' by UX practitioners and determine ethical concerns based on this list. The figures provided aimed to demonstrate the variety of Dark Patterns and their problematic nature.

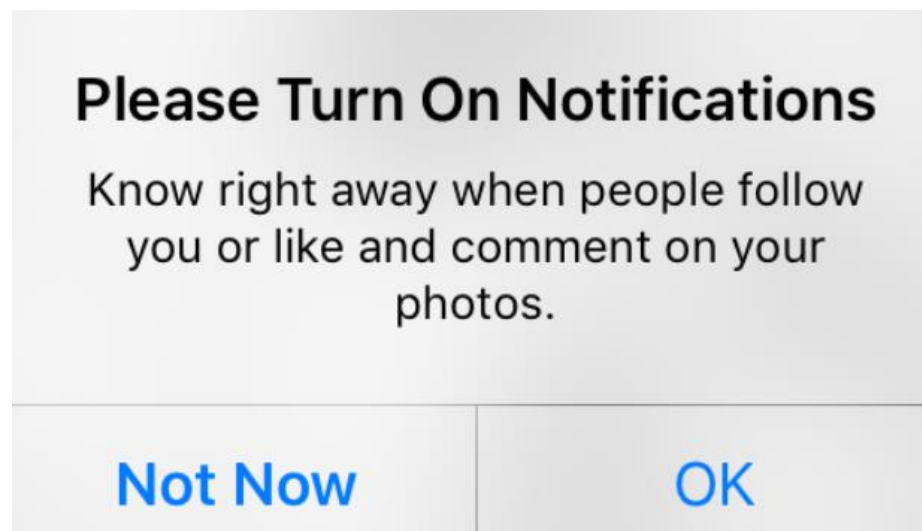


Figure 1. An example of the 'nagging' dark pattern from Instagram. This notification offers no way to permanently dismiss and will continually pop-up or 'nag' *ad infinitum* (Gray et al., 2018, p. 6)



Figure 2. An example of an ‘obstruction’ or ‘labyrinthian navigation’ (depending on which name the author chooses) on iOS 6. The option to limit ad tracking is buried in a confusing and labyrinthian menu structure (Gray et al., 2018, p. 7).

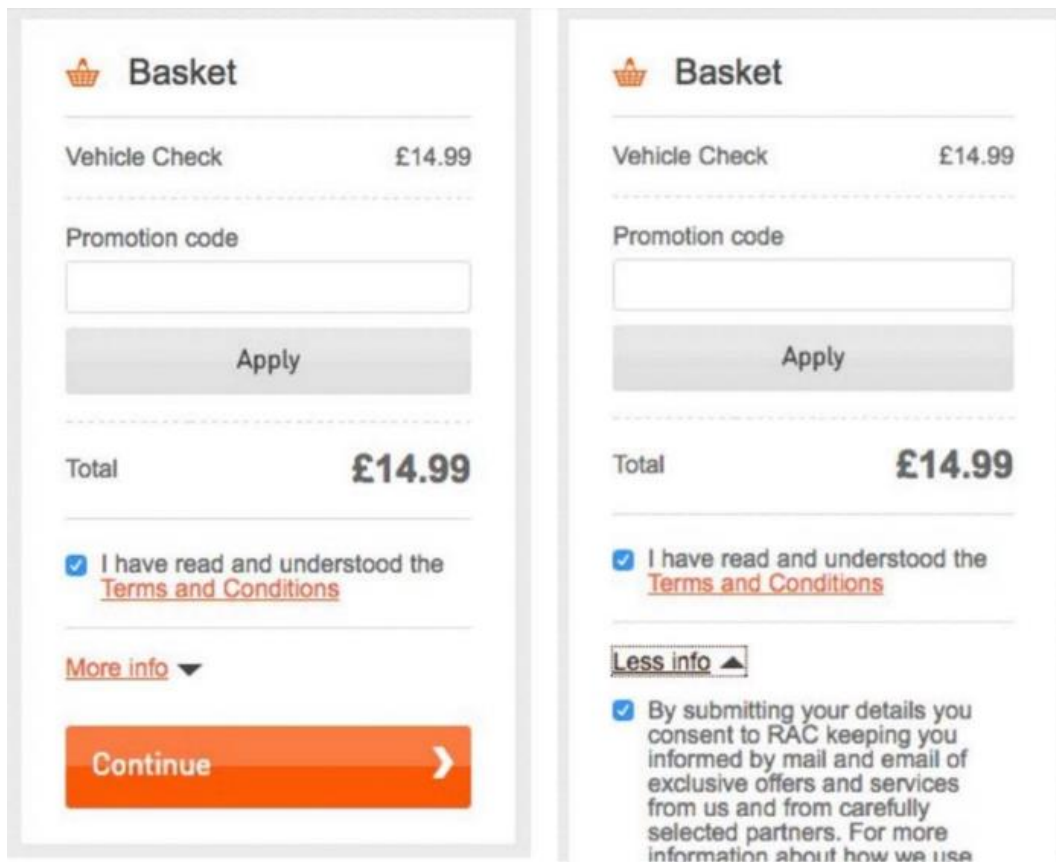


Figure 3. An example of two Dark Patterns working together. 1. ‘Preselection’ which automatically selects an option beneficial to the platform. 2. ‘Hidden information’ which visually obscures important information (Gray et al., 2018, p. 8).

Examples from Luguri & Strahilevitz (2021)

The aim of Luguri & Strahilevitz's (2021) research was to demonstrate the effectiveness of Dark Patterns. Furthermore, they propose legal policies to discourage the use of Dark Patterns. The figures provided aimed to demonstrate the effectiveness of Dark Patterns and generate evidence for their legal advice.

Membership Status

Canceling your membership?

Are you sure you want to cancel your membership? You will no longer receive membership pricing on all our products.

CONTINUE

CANCEL

Figure 4. An example of the Dark Pattern 'trick question' found on the Pressed Juicery's cancellation page. It is unclear if the CONTINUE button continues with the cancellation process or continues the membership. This lack of clarity is compounded by the CANCEL button which makes it unclear if it confirms a request to cancel or cancels the cancellation request (Luguri & Strahilevitz, 2021, p. 49).



Figure 5. An example of 'forced continuity' combined with 'roach motel' found on a Sony PlayStation. There is no way to purchase the product without the subscription auto renewing. The subscription's auto-renew feature can only be disabled after purchase by digging into a settings menu that is neither simple nor intuitive (Luguri & Strahilevitz, 2021, p. 56).

Examples from Mildner et al. (2023)

The aim of Mildner et al.'s (2023) research was to fill a gap they identified in the research of Dark Patterns. They sought to identify instances of Dark Patterns that could help to explain a perceived “lack of users’ self-determination regarding control over personal data and time spent on [social networking services]” (Mildner et al., 2023, p. 1). The figures provided aimed to demonstrate this lack of self-determination and control.

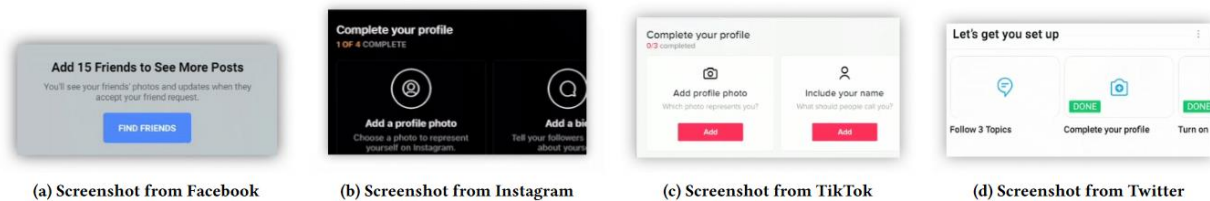


Figure 6. Four examples of ‘gamification’ found on various social media platforms. Performing the requested tasks will connect the user to followers, friends, etc. Mildner et al. (2023) argue that this may be motivated by the platform’s desire to keep people socially engaged on the platform for longer than they may intend and encourage users to share their information with more people than they may be comfortable with (Mildner et al., 2023, p. 9).

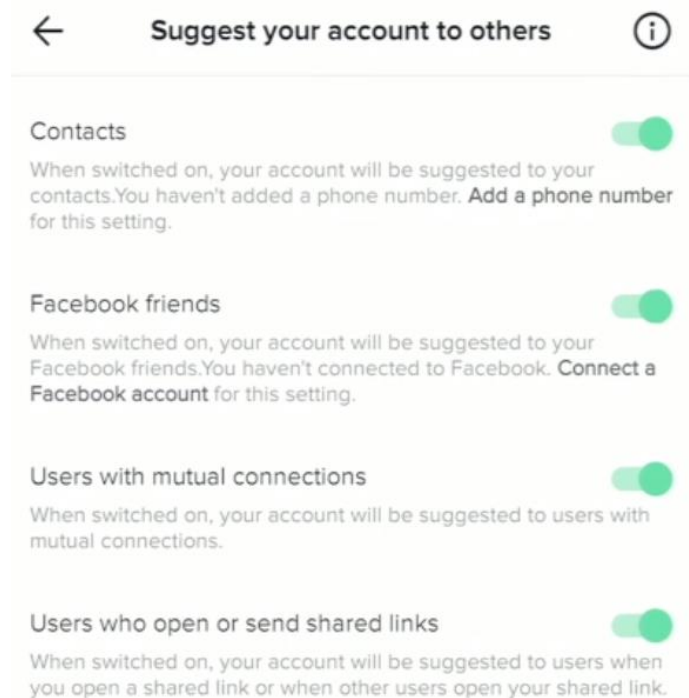


Figure 7. An example of ‘social brokering’ working in tandem with ‘preselection’ and ‘privacy zuckering’ found on TikTok. These options are on by default and connect with people both on and off the platform using personal information to do so in some cases (Mildner et al., 2023, p. 10).

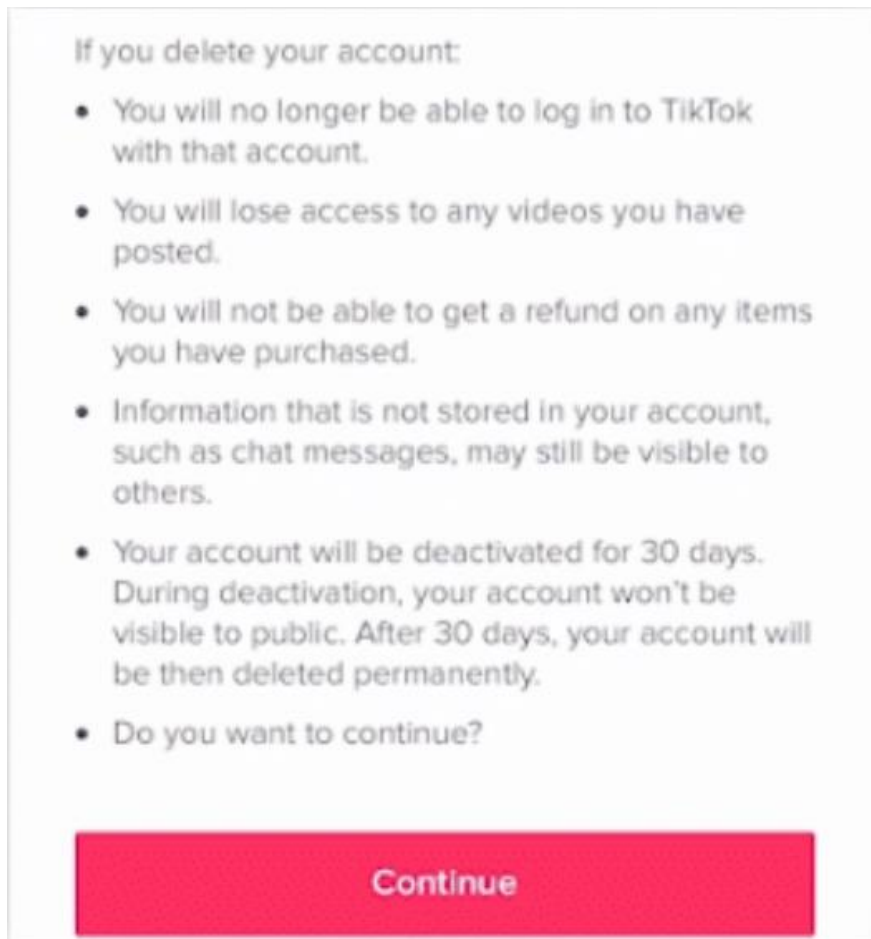


Figure 8. An example of a 'forced grace period' found on TikTok. There is no way to immediately delete an account, users must wait for 30 days for their account to be deleted (Mildner et al., 2023, p. 11).

An example from Nouwens et al. (2020)

The aim of Nouwens et al's (2020) research was to demonstrate how Dark Patterns could be used to undermine the GDPR and how effective these interfaces might be in increasing rates of consent. The figures provided demonstrate how consent management platforms can use Dark Patterns to potentially undermine the GDPR. All three images provided in this section are from the same consent form and progress in the order they can be found.



We value your privacy

Our site is supported by advertising and we and our partners use technology such as cookies on our site to personalize content and ads, provide social media features, and analyze our traffic. Click "I Accept" below to consent to the use of this technology across the web. You can change your mind and change your consent choices at any time by returning to this site and clicking the Privacy Choices link.

By choosing I Accept below you are also helping to support our site and improve your browsing experience.

[More Options](#) | [See Vendors](#)

Powered by **Quantcast**

Figure 9. An example of ‘false hierarchy’ found on <https://sourceforge.net/> The highlighting of “I ACCEPT” is considered to indicate it as a preferred option in UX design, though no meaningful reason is presented to prefer it (Nouwens et al., 2020, p. 3).

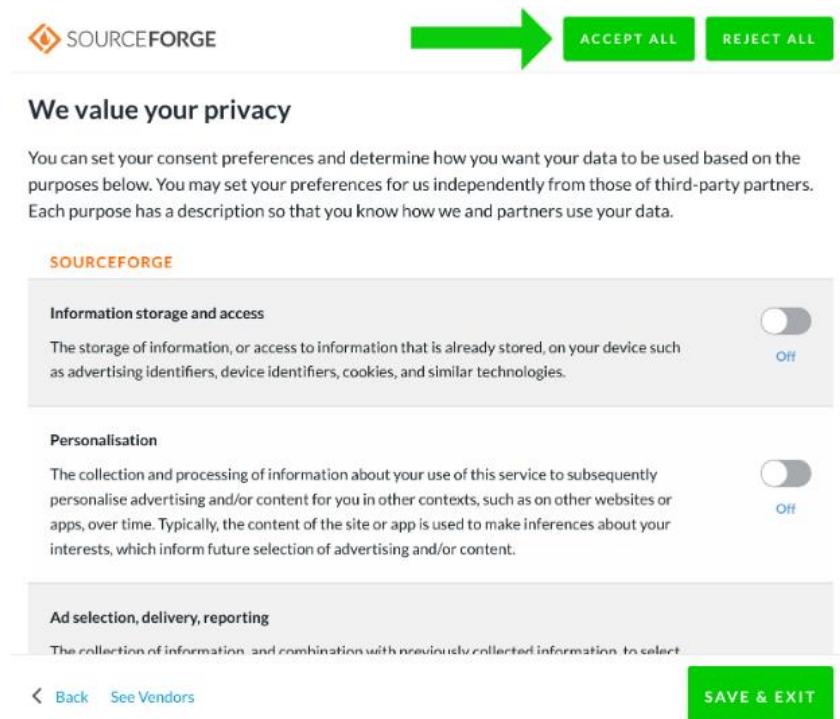


Figure 10. An example of ‘forced dialogue interaction’ that only comes up when the above “I DO NOT ACCEPT” button is clicked. Screens like this may forgo the “REJECT ALL” button at the top and preselect each category to be toggled on, making the denial of consent more difficult (Nouwens et al., 2020, p. 3).



ACCEPT ALL

REJECT ALL

We value your privacy

You can set consent preferences for each individual third-party company below. Expand each company list item to see what purposes they use data for to help make your choices. In some cases, companies may disclose that they use your data without asking for your consent, based on their legitimate interests. You can click on their privacy policies for more information and to opt out.

Google: If you accept all purposes, Google ads will be personalized. [Google's Privacy Policy](#).

COMPANY	OFF/ON
1020, Inc. dba Placecast and Ericsson Emodo	<input type="checkbox"/> 
1plusX AG	<input type="checkbox"/> 
2KDirect, Inc. (dba iPromote)	<input type="checkbox"/> 
33Across	<input type="checkbox"/> 

[Back to Purposes](#)

SAVE & EXIT

Figure 11. An additional 'force dialogue interaction' that again comes up when consent is not given in figure 9. These screenshots may also be an example of 'labyrinthian navigation' as these menus may be confusing to end users (Nouwens et al., 2020, p. 3).

Appendix B - Alternative Tables

This appendix provides two alternative tables, both provided by Mildner et al. (2023), that comprehensively summarize the various Dark Patterns found in the research. I chose to use the more succinct table provided by Luguri & Strahilevitz (2021) within the body of my thesis due to its descriptiveness and discernment, but there are some advantages of Mildner et al's (2023) tables and so they are worth adding to this thesis here.

Table 2 provides an entire list of all the discussed Dark Patterns in the extensive literature Mildner et al. (2023) reviewed. This table can be helpful to 1) see the full list of Dark Patterns as described in the research and 2) identify discrepancies between authors who may be referencing a similar or even identical Dark Pattern but using slightly different terminology depending on their sources. No description is provided for these Dark Patterns, but the names are, for the most part, fairly descriptive of what they seek to accomplish.

Table 3 provides Mildner et al's (2023) additions based on the inductive codes they determined using the questionable and problematic interfaces they examined that did not fit into the list provided in Table 2. The strategies they examined as related to the Dark Patterns found in Table 3 are very relevant to the content of this thesis. Furthermore, the discussion of these strategies were of a high research quality. However, they give the final Dark Pattern coded in this table the name "plain evil" which they describe as "interface elements suddenly change functionalities or are being exchanged for alternative ones" (Mildner et al., 2023, p. 15). I believe this term is problematic and their description fails to capture the extremity of 'evil'. As a result, this table was not used within the body of this thesis.

Author	Dark Pattern	F	I	Ti	Tw
Brignull [7]	Bait And Switch	●	○	○	●
	Confirmshaming	●	●	●	●
	Disguised Ads	○	●	○	●
	Forced Continuity	○	○	○	○
	Friend Spam	○	○	○	○
	Hidden Costs	○	○	○	○
	Misdirection	●	●	●	○
	Price Comparison Prevention	○	○	○	○
	Privacy Zuckering	●	●	●	●
	Roach Motel	●	●	●	●
	Sneak Into Basket	○	○	○	○
	Trick Question	○	○	○	○
	Conti & Sobiesk [9]	Coercion	○	○	○
Confusion		●	○	○	●
Distraction		●	●	●	●
Exploiting Errors		○	○	○	○
Forced Work		●	●	●	●
Interruption		●	●	●	●
Manipulating Navigation		●	●	●	●
Obfuscation		●	●	●	●
Restricting Functionalities		●	●	○	○
Shock		○	●	○	○
Trick	○	○	○	○	
Zagal et al. [47]	Grinding	○	○	○	○
	Impersonation	○	○	○	○
	Monetized Rivalries	○	○	○	○
	Pay To Skip	○	○	○	○
	Playing By Appointment	○	○	○	○
	Pre-Defined Content	○	○	○	○
	Social Pyramid Schemes	○	○	○	○
	Attention Grabber	●	●	●	○
Greenberg et al. [20]	Bait And Switch	○	○	○	○
	Captive Audience	○	○	○	○
	Disguised Data Collection	○	○	○	○
	Making Personal Info. Public	○	○	○	○
	The Milk Factor	○	○	○	●
	Unintended Relationships	○	○	○	○
	We Never Forget	○	○	○	○
	Legend:	F - Facebook			
	I - Instagram				
	Ti - TikTok				
	Tw - Twitter				
Bösch et al. [5]	Address Book Leeching	●	●	●	●
	Bad Defaults	●	●	●	●
	Forced Registration	○	○	○	○
	Hidden Legalese Stipulations	●	●	●	●
	Immortal Accounts	○	○	○	○
	Information Milking	●	○	○	○
	Privacy Zuckering	●	○	●	○
	Shadow User Profiles	○	○	○	○
	Forced Action	●	●	●	●
	<i>Gamification</i>	●	●	●	●
	<i>Social Pyramid</i>	●	●	●	●
	Interface Interference	●	●	●	●
	<i>Aesthetic Manipulation</i>	●	●	●	●
	<i>False Hierarchy</i>	●	●	●	●
<i>Hidden Information</i>	●	●	●	●	
<i>Preselection</i>	●	●	●	●	
<i>Toying With Emotions</i>	●	●	●	●	
Nagging	●	●	●	●	
Obstruction	●	●	●	●	
<i>Intermediate Currency</i>	●	●	●	●	
Sneaking	●	●	○	○	
Gray et al. [19]	Automating The User	●	●	○	●
	Controlling	●	●	●	●
	Entrapping	○	○	○	○
	Misrepresenting	●	●	●	●
	Nickling-And-Diming	○	○	○	○
	Two Faced	○	○	○	○
	Forced Action (see Gray et al. [19])				
	<i>Forced Enrollment</i>	○	○	○	○
	Misdirection	●	●	●	●
	<i>Pressured Selling</i>	●	●	●	●
<i>Visual Interference</i>	●	●	●	●	
Gray et al. [18]	Obstruction (see Gray et al. [19])				
	<i>Hard To Cancel</i>	●	●	●	●
	Scarcity	●	●	●	●
	<i>High-Demand Messages</i>	○	○	○	○
	<i>Low-Stock Messages</i>	○	○	○	○
	Sneaking (see Gray et al. [19])				
	<i>Hidden Subscriptions</i>	○	○	○	○
	Social Proof	●	●	●	●
	<i>Activity Notifications</i>	○	○	○	○
	<i>Testimonials</i>	○	○	○	○
Mathur et al. [31]	Urgency	●	●	●	●
	<i>Countdown Timer</i>	○	○	○	○
	<i>Limited-Time Messages</i>	○	○	○	○

Table 2. List of Dark Patterns and the platforms that use them by Mildner et al. (2023, p. 7)

Strategy	Theme	No.	Inductive Code	F	I	Ti	Tw
Engaging Strategies	Interactive Hook	1.	Addictive Design	○	●	●	●
		2.	Autoplay Content	●	●	●	●
		3.	Fear Of Missing Out	○	○	○	●
		4.	Gamification	●	●	●	●
		5.	Infinite Scrolling	●	●	●	●
		6.	Pull To Refresh	●	●	●	●
		7.	Reduced Friction	●	●	●	●
	Social Brokering	8.	False Content Customisation	●	○	○	○
		9.	Regression Toward The Mean	●	○	●	●
		10.	Social Connector	●	●	●	○
Governing Strategies	Decision Uncertainty	11.	Decision Uncertainty	●	○	○	●
		12.	Clinging To Accounts	●	●	●	●
		13.	Persuasive Language	●	●	●	●
	Labyrinthine Navigation	14.	External Solution Search	○	●	○	○
		15.	Labyrinth	●	●	○	●
		16.	Hidden In Plain Sight	●	●	●	●
	Redirective Condition	17.	Auto Accept Third Party Terms	○	○	○	●
		18.	Decision Governing	●	●	○	○
		19.	Forced Access Granting	○	○	●	○
		20.	Forced Dialogue Interaction	●	○	●	○
		21.	Forced Grace Period	●	●	●	●
		22.	Plain Evil	○	○	●	○
Total				16	14	15	15

Table 3. Mildner et al's (2023, p. 9) additional Dark Patterns and the platforms that use them. F = Facebook, I = Instagram, Ti = TikTok, & Tw = Twitter (i.e., X)