



Underlagsrapport 2021:1

Barriers to a well-functioning digital market

Effects of visual design and information disclosures on consumer detriment

Underlagsrapport 2021:1 *Barriers to a well-functioning digital market – effects of visual design and information disclosures on consumer detriment*

Konsumentverket 2021

I en underlagsrapport är resultat, slutsatser och förslag som förs fram författarens egna.

Preface

This research was commissioned by the Swedish Consumer Agency and was carried out by Karlstad University, Service Research Center (CTF) in between August – December 2020. The purpose of this report is to provide in-depth information for the Swedish Consumer Agency on barriers to a well-functioning digital market. In particular, to present an overview of user interface designs that affect consumers' detriment, constrain their autonomy of decision-making, and inform the Swedish Consumer Agency about measures that might help to overcome malicious designs.

We would like to thank all interview respondents and the Swedish Consumer Agency for their commitment and for providing us with input and feedback on our results.

Karlstad, 2021

Agnieszka Kitkowska, CTF, Karlstad University
Erik Wästlund, CTF, Karlstad University

Abstract

With growing access to the internet globally, online shopping is on the rise and the number of digital consumers is increasing every year. This trend seems to persist, and it is predicted to accelerate further in the future. However, this growth is associated with certain risks to the well-being of both consumers and markets. One such risk relates to inefficient information disclosures and tricky user interface designs in digital applications that may cause consumer detriment and, consequently, negatively impact the digital market. In this research, we attempted to shed light on two issues that might prompt such adverse effects—dark patterns and incomprehensive terms and conditions.

This report intends to identify barriers to the well-functioning digital market through a review of existing research and a qualitative empirical study (expert interviews) and provide guidance on overcoming such obstacles. Further, by applying a behavior change framework, we identified what measures could be taken, mainly by regulators, to ensure a well-functioning market.

A systematic literature review was conducted to identify dark patterns in digital markets. Based on the findings, we re-categorized the identified dark patterns according to their impact on consumers and mapped them to the behavior change COM-B (Capability, Opportunity, Motivation-Behavior) model and behavioral change wheel. Additionally, our literature review and a series of expert interviews were used to recognize barriers to understanding and acting according to the disclosures included in terms & conditions agreements.

Based on our findings, we present a discussion about the concept of consumer vulnerability, which is critical to understanding consumer detriment. Moreover, our thematic analysis of the expert interviews defines the obstacles in the current policies and intervention functions that might indirectly and directly negatively influence consumers' choices. We summarize and discuss our findings by presenting a list of barriers and recommendations to overcome them, as proposed by our experts. These research findings might be useful to policymakers and regulatory bodies and could be utilized as a basis for future research on this topic.

Table of contents

List of abbreviations	6
Definitions	7
1 Introduction	8
2 Background	10
2.1 Determinants of Behavior	10
2.2 Human-Computer Interaction and Decision-Making	11
2.3 Behavior Change	16
2.4 Objectives.....	17
3 Methods.....	19
3.1 Systematic Literature Review.....	19
3.2 Expert Interviews.....	21
4 Findings.....	22
4.1 Dark Patterns.....	22
4.2 Terms & Conditions	38
4.3 Consumer Vulnerabilities.....	43
4.4 Measures to Tackle Barriers	45
5 Summary and Discussion	49
5.1 Dark Patterns.....	49
5.2 Terms & Conditions	50
5.3 A Well-functioning Digital Market.....	51
5.4 Ethical issues	54
5.5 Effects of Dark Patterns on Companies	56
6 Bibliography.....	58
Appendix.....	1

List of abbreviations

COM-B	Capability, Opportunity, Motivation - Behavior
EAST	Make it Easy, Attractive, Social, Timely
EU	European Union
EULA	End-User License Agreements
GDPR	General Data Protection Regulation
HCI	Human-Computer Interaction
HTML	Hypertext Markup Language
MIND SPACE	Messenger, Incentives, Norms, Defaults, Salience, Priming, Affect, Commitment, and Ego
OECD	Organization for Economic Co-operation and Development
T&Cs	Terms and Conditions
UI	User Interface
UX	User Experience

Definitions

Bias. A systematic pattern of deviation from norm or rationality in judgment and decision making. Individuals perceive an input based on their “subjective” rationality, and such perception dictates how they behave.

Heuristic. An experience-based strategy for solving a problem or making a decision often provides an efficient means of finding an answer but cannot guarantee a correct outcome.

Human-Computer Interaction. This discipline is concerned with the design, evaluation, and implementation of interactive computing systems for human use and the study of major phenomena surrounding them. It is an interdisciplinary field that incorporates computing, cognitive sciences, and human factors engineering.

User Interface. This term refers to any interaction with a system, including physical, perceptual, and conceptual interactions.

Choice architecture. This refers to the physical and symbolic environment that faces decision-makers at the point when they make a decision. Such an environment can have a significant impact on the choices and potentially make them more predictable.

Online choice architecture. The principles of choice architecture are applied in the digital context, mostly through the UI design, at three levels: physical, perceptual, and conceptual.

1 Introduction

A well-functioning economy is one that safeguards fair market competition and the availability of information for all involved parties. This definition also applies to the digital economy, where the availability of information is often imbalanced and, in some cases, affects fair competition. In a similar vein, a well-functioning market should enable consumers to make the most optimal choice and guarantee that they understand the consequences of their transactions, their rights, and/or obligations that such transactions may impose. It should also allow them to make autonomous decisions in a way that does not limit their ability to choose between different services or product providers.

In the context of the digital space, information imbalance might be caused by a complete lack of information or information presented incomprehensibly. For instance, transaction information might be either cognitively (e.g., through the use of difficult language) or visually (e.g., through the concealment of information in small font) inaccessible to consumers. These deficiencies in information may result in consumer detriment. As defined by the EU, consumer detriment occurs when “market outcomes fall short of their potential, resulting in welfare losses for consumers.” Consumer detriment can have two forms. The first is structural detriment, which can be explained as a loss of aggregate consumer welfare because of market or regulatory failure. The second one is personal detriment, which reflects a difference between the expected value of the good or service and how it differs from the actual value that the consumer receives.

Consumer detriment may be either revealed or unrevealed (hidden). In the case of the former, consumers are aware of the detriment. However, the latter is hidden and falls within the definition of structural detriment because it affects aggregate consumers. In this report, we are mostly concerned with personal detriment and, to some extent, both revealed and unrevealed personal detriment. To this end, we adopt the definition of personal detriment proposed by the Organization for Economic Co-operation and Development (OECD): “negative personal outcomes for individual consumers that they became aware of following the purchase or use of a good or service, measured relative to what would reasonably have been expected given the type of transaction” [1].

Unbalanced information about market transactions is one of the factors that may lead to an increase in consumer detriment. Hence, it is essential to gain an in-depth understanding of factors that cause an increase in information disparity. In the digital market context, it is a system’s design that presents consumers with information. Mostly, all the information required for a transaction is obtainable through the user interface (UI) (here, we refer to both functional and visual aspects of the interface, including online choice architecture). Hence, the UI design is important for understanding information disparity and the associated consumer detriment, and research on barriers integrated within the UI design could help shed light on the barriers to a well-functioning digital market.

This report focuses on consumer detriment through an overview of barriers to a well-functioning market. In particular, this research investigates dark patterns and barriers to understanding and acting on terms and conditions (T&Cs). Further, the current report describes a series of measures that could be applied to reduce consumer detriment by changing companies' behavior via modifications to intervention and policy design.

2 Background

In many countries, legislation requires online companies to disclose certain information to ensure that consumers can make informed decisions. Therefore, in some ways, online disclosures are consumers' guides that, with appropriately displayed information, should maintain consumers' engagement in the digital economy and the market's efficient functioning [2]. Thus, the well-functioning market relies on information disclosures:

“Information about price, quality, and attributes allow buyers to make the best use of their budget by finding the product whose mix of price and quality they most prefer. In turn, buyers' ability to locate preferred products gives sellers an incentive to compete to improve their offerings by allowing buyers to find and reward the seller (with patronage) whose offer they prefer. Without such information, the incentive to compete on price and quality will be weakened, and consumer welfare will be reduced” [3].

To make an informed decision regarding a transaction, consumers must be provided with adequate information [2]. Four key elements contribute to such information. First, consumers should be informed about the attributes related to the price of the product or service. Second, they should be provided with the T&Cs of the sale. Third, there should be sufficient information about payment and delivery. Fourth, consumers should also be provided with information about their rights after the purchase. Such information disclosures may benefit markets, as they do not interfere with free trade and ensure that consumers are left with free choices about the goods or services. Additionally, disclosures help to ensure consumer autonomy. However, particularly in the online context, disclosures may be problematic, for instance, due to information asymmetry between the consumer and service provider. Further, consumers may experience information overload, which might be detrimental. For example, consumers may resign from purchasing products that could be financially beneficial. Additionally, the information presented to consumers might be complicated and require time to comprehend, especially considering the display of T&Cs or privacy policies.

2.1 Determinants of Behavior

Interaction with any technology is dependent on the technology's design. Here, design refers to all aspects of technology, including the physical, perceptual, and conceptual aspects of technology. In the context of digital markets, the most prominent elements that may affect consumers' behavior are the UI, particularly its visual aspects (e.g., layout, colors, and prominence of different features) and the represented content (e.g., text and images). All UI designs guide users to make specific decisions, and it may be impossible to create an entirely neutral interface that does not affect the steps that a user takes to complete their action. Unfortunately, in some cases, companies abuse this feature of UI designs to influence users' decisions, for instance, by manipulating consumers' purchasing choices or exploiting the psychological vulnerabilities of human nature.

It is commonly recognized that autonomy should be granted to consumers in digital markets to enable their most optimal and informed choices. However, companies are known to use certain tactics that prevent consumers from learning information

that might be useful. One of the tactics that affect consumers' well-being identified in past research is dark patterns. Dark patterns are defined as UIs that are purposefully designed to confuse users. As a result, users cannot follow their preferences and desires and become subject to manipulations [4]. The second tactic is the use of T&Cs to convey critical information, which, in the current UIs, is often disregarded or perceived as unimportant. T&Cs, also referred to as Terms of Service (ToS), are usually non-negotiable agreements that consumers must approve to use a service. Currently, most online companies present T&Cs as merely marginal information; yet, it may contain disclosures important to consumer well-being, such as rights after purchase and additional costs, and more. T&Cs agreements are a subject of debate because they frequently cause an imbalance wherein "one party is generally less powerful in terms of access to information and resources" [5]. Thus, although T&Cs are often overlooked by consumers due to their secondary role (as the consumer's primary task is to complete a purchase) during online interaction or a design that diminishes their importance (e.g., fine print and hyperlinked text), they contain information that might be of significance for the consumer. The importance of information included in T&Cs could be summarized as follows:

"this 'fine print' now exists within the corners and margins of nearly all of our online activities and often contains clauses that govern several important aspects of our lives, including copyright and ownership policies, dispute and jurisdiction information, acceptable use, even labour terms in many contexts" [5].

In the next section, this paper presents an overview of human-computer interaction (HCI) and decision-making processes to understand better the origins and workings of dark patterns and the role of information presented in UIs. It also explains the behavior change framework, which was employed in the current work to identify measures needed to change consumers' and companies' behavior.

2.2 Human-Computer Interaction and Decision-Making

Making decisions lies at the core of HCI. This function is essential when people interact with technologies as consumers in e-commerce settings. Consumers may know what they would like to purchase, that is, the goods or services that best suit their current needs or fulfill their desires. However, at times, consumers do not align their choices with their original desires or preferences. Such a gap between consumers' attitudes and behaviors is not uncommon. In psychology and behavioral sciences, such phenomena have been recognized as being typical of human behavior. It is frequently explained with dual-process theories about the cognitive processing that takes place during decision-making. According to the dual-process approach, people make decisions using the two different information processing modes that work in parallel. The first mode is automatic, fast, and uncomplicated, and is responsible for decisions that do not require working memory [6][7]. The second mode involves more complex, slower information processing and is often considered more rational (in an economic sense) [7]; this mode is reflective and requires working memory [6]. Most people's daily decisions rely on automatic processing because it is quicker and requires less cognitive effort. Such information processing uses mental shortcuts—heuristics—that enable people to

solve problems and make judgments quickly and efficiently. However, such heuristic information processing may be prone to cognitive biases (e.g., subjective patterns and inclinations towards choices that are not necessarily the most optimal) and place people at a disadvantage.

In many studies, the automatic thinking mode was shown to guide the majority of decisions. These results were utilized by choice architects, who exploited different psychological effects (biases and heuristics) associated with intuitive thinking through the choice architecture design. One of the strongest promoters of this approach is the Nobel laureate Richard Thaler, who, together with Cass Sunstein, described how biases and heuristics could be applied in choice architectures to improve human well-being. This led to the development of the concept of nudging. A nudge is defined as “any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives” (p.6, Thaler & Sunstein, 2008). Since the development of nudges, governments have applied them to improve people’s well-being, and some of these applications have been successful. For instance, the recognition heuristic was shown to be reliable in predicting consumer behavior. In the context of healthcare, fast-and-frugal heuristics help physicians make diagnoses. Further, in some countries, the use of the default heuristic leads to higher organ donations and results in shorter waiting times and a lower number of deaths [9].

Even though by its original definition, the aim of nudging is to help people improve their well-being, nudging might be used maliciously. Dark patterns can be classified as a malicious form of nudging. As explained earlier in this paper, dark patterns are parts of UI design that trick people into specific actions. Designers of dark patterns utilize their knowledge about human behavior (e.g., psychology, behavioral science, and nudges) to develop systems or services that deceive users and result in decisions with less beneficial outcomes for the user [10]. Therefore, both nudges and some dark patterns seem to have a common basis—biases and heuristics triggered during decision-making.

2.2.1 Nudging

With the growing reliance on technology, the concept of nudging has been adopted in digital systems. It appears that online decisions are also based on biases and heuristics that may drive people’s online choices. Hence, digital nudges can be defined as nudges that alter users’ preferences at the point of interaction [8]. Here, the slight adjustments in either the content or visual display of a UI may lead to users’ action changes.

Digital nudges can be applied in UI design and affect users’ choices by exploiting particular psychological vulnerabilities [8]. For example, in a UI that contains pre-selected boxes, the status quo bias is used in so-called binary choice decisions. Additionally, specific positioning of UI elements containing choice options, such as choosing between two products, may trigger primacy and recency effects. Other heuristics that could be activated with UI designs are the middle-option bias, anchoring (e.g., slider endpoints), social norms (e.g., display of popularity and honesty codes), and loss aversion (e.g., limited availability) (see Table 1 for definitions).

Various psychological vulnerabilities are also utilized in persuasive technologies (i.e., any interactive systems designed to change people's attitudes and behaviors [11]). Researchers have identified digital nudging elements in such technologies and have emphasized the difference between nudging and persuasive design: The main difference is that persuasive technology does not predict people's behavior and does not use coercion or deception [9], although it builds on similar psychological premises, such as anchoring, customized information, decision staging, default settings, framing, informing, limited time window, praise and reward (gamification), priming, reminders simplification (reduction), social influence, and warning.

The concept of nudging has also been studied in the context of HCI [12], and 23 mechanisms of nudging developed in HCI have been identified and clustered into six categories.

- a. Nudges that *facilitate* decision making by reducing cognitive effort (through exploiting the status quo bias)
- b. *Confronting* nudges that aim to pause action by eliciting a doubt (by using the regret aversion bias)
- c. *Deceiving* designs (that exploit, e.g., a decoy effect)
- d. *Social influences*, which exploit human nature through the effects of conformity and social expectations
- e. *Fear* nudges that elicit emotional feelings of loss, fear, or uncertainty
- f. *Reinforcement* nudges that aim to reinforce behaviors through their continued presence in an individual's thinking

HCI nudges can also be categorized based on whether they trigger the automatic thinking mode or reflective thinking mode [12], as shown below.

1. Interventions that are transparent and facilitate consistent choice (by triggering the reflective thinking mode), e.g., comparison of social norms, such as energy consumption, that nudges consumers towards energy-saving choices
2. Interventions that are transparent and influence behavior (by triggering the automatic thinking mode), e.g., changes to the default options in printing settings on the computer that encourage people to print on both sides of a page to reduce paper waste
3. Non-transparent interventions that manipulate choice (by triggering the reflective thinking mode), e.g., the addition of irrelevant choices to a set of options to increase the value of a particular choice
4. Non-transparent manipulations of behavior (by triggering the automatic thinking mode), e.g., opt-in policies in a system that asks users to schedule the next appointment for vaccination and automatically sets up reoccurring appointments

Researchers have also identified 20 psychological effects triggered by cues embedded in UI design that change people's attitudes and behaviors [12]. Although these findings are limited to the research on privacy-related decision-making and show how specific visual cues may either deteriorate or enhance privacy attitudes and behaviors, privacy-deteriorating nudges might be applicable to the context of e-commerce and consumers.

In sum, nudges and their psychological underpinnings can be misused by online companies. It is important to note that nudges may backfire and negatively affect users [13]. Additionally, they might be intrusive and limit people's autonomy. Thus, when misused or purposefully implemented in a hostile manner, a nudge may quickly become a dark pattern that no longer serves the purpose of improving well-being. Considering this close relationship between nudges and dark patterns, we believe that understanding the different psychological aspects of decision-making processes is crucial to help designers identify when their designs may become *dark*.

Based on the findings from research on HCI and decision making, Table 1 presents select psychological biases and heuristics related to the current work. This list of biases and heuristics is not exhaustive, but it informs the latter sections of the current paper.

Table 1: Heuristics and biases frequently used to nudge. This list is not complete.

Name	Definition
Affect heuristic	“People make judgments based on the representations of objects and events in their minds that are tagged to vary in degrees with affect” (“goodness” or “badness” experienced as a feeling or demarcating a positive or negative quality of stimulus) [14]. For example, images showing babies tend to result in more positive affective states and might lower the perception of risks that a current activity may carry.
Anchoring	Under uncertainty (e.g., a lack of information), the decisions are biased towards the starting point used to calculate estimates [15]. For example, the initial price for a product presented before the purchasing decision will set an anchor point, and prices lower than the initial cost will seem more reasonable.
Availability heuristic	This is defined as “an assessment of accessibility in which frequencies or probabilities are judged by the ease with which instances come to mind” [7]. For example, after viewing several adverts on the luxurious lives of people who won a lottery, one could mistakenly think that their chances of winning are higher than they are.
Bandwagon effect	People tend to infer that particular behavior is right when a lot of people do it and tend to follow that behavior [16]. For example, when many people are buying a particular product, others are more likely inclined to purchase it.
Decoupling	Evaluation of costs and benefits may differ according to the form of payment (because it may be decoupled from consumption)[17]. For example, retailers offering to delay payment for a product aim to decouple the purchase from payment to increase purchasing possibility.
Default effect	This is defined as a preference for the default option over changing it or acceptance of the status quo [18]. For example, when completing a purchase, one of the delivery options might be pre-selected. The consumer may overlook other options, leaving the delivery as is.

Name	Definition
Endowment effect	This is “the tendency when people demand much more to give up an object than they would be willing to pay to acquire it” [19]. For example, when a retailer offers something free attached to the purchase, the consumer is immediately endowed with a product and less likely to give up the purchase.
Framing	The decision frame can be designed in many ways that control the decision problem’s presentation, thus influencing the final decision [17]. For example, a small font can be used to make disadvantageous information about the product less visible to the consumer and more likely to be overlooked.
Hyperbolic discounting	Individuals behave inconsistently over time and tend to value present, smaller rewards more over the future, larger ones [20]. For example, people would rather receive a SEK100 discount now than a SEK150 discount next month.
Information asymmetry	This is “acquired from the agency theory developed to address the principal-agent problem, i.e., the difficulties that arise under conditions of incomplete and asymmetric information when a principal hires an agent to pursue the principal’s interests” [12]. For example, a used car salesperson might have more information about the reliability of a given car than potential buyers and can use this information to make the car seem better than it actually is.
Instant gratification	This heuristic is characterized by sacrificing the future to gain immediate pleasure/satisfaction. In this trade-off, the reward is quick, and the cost is delayed [21]. For example, consumers will be more likely to select same-day delivery instead of waiting for weeks.
Image motivation	The desire to be positively perceived by others is a driver of pro-social behavior [22]. For example, a consumer might be encouraged if charitable donations are added to a purchase and consumers that contributed to the charity are named on a public list of donors.
Loss aversion	“The dis-utility of giving up an object is greater than the utility associated with acquiring it” [19]. For example, a consumer may focus on one investment that has lost money while ignoring other investments.
Messenger effect	“The weight given to information depends on the automatic reactions to the perceived authority of the source of that information” [23]. For example, reviews of a product by an individual labeled as an “expert” might have a greater impact on consumer choice than reviews without such labels.
Mere exposure	Exposure to something changes how people perceive it; that is, people are less uncertain about it. For example, some companies may invest only in products that consumers are familiar with, as they have been exposed to the brand name several times.
Representativeness	“Some probability judgments (the likelihood that X is a Y) are mediated by assessments of resemblance (the degree to which X ‘looks like’ a Y)” [24]. For example, ads of a specific car that always present a happy family are more likely to make people conclude that such cars are only good for families.

Name	Definition
Scarcity bias	This is characterized by the “tendency to attribute more value to an object because we believe it will be more difficult to acquire in the future” [13]. For example, companies may use banners showing consumers that a given product is only available at a discounted price for another 24 hours.
Social norms	Unwritten rules are standards that are understood by members of a group that guide and/or constrain behavior [17]. For example, a water supply company may present consumers with their neighbors’ water consumption - their neighbors manage their water consumption more carefully and save money. Thus, influenced by such information, consumers’ behavior may change.
Status quo	“Humans tend to prefer options that cause no change in their state and/or require no action on their part” [25]. For example, sending consumers reinforcing messages about one brand may increase customer loyalty and increase their intention to renew contracts.

2.3 Behavior Change

Some aspects of UI designs, such as dark patterns or ineffective disclosures, might negatively affect consumers. Consumers influenced by such designs behave in a specific and predictable way, frequently not realizing that the system might exploit their vulnerabilities and trick them into a particular choice.

One possible way of helping consumers to make more informed choices is to apply certain behavior interventions. However, it remains unclear what interventions should be used in the context of digital markets. Various frameworks employed in psychology and behavioral sciences might help to identify such interventions, for instance, MIND-SPACE (Messenger, Incentives, Norms, Defaults, Salience, Priming, Affect, Commitment, and Ego) and the framework of the Cochrane Effective Practice and Organisation of Care Review Group. Another framework that could be used as a tool by policymakers to connect behavioral knowledge with policy problems is EAST (Make it Easy, Attractive, Social, Timely) [26]. However, this framework is mostly concerned with techniques used to present information, mainly the information from privacy notices.

In this work, we consider a different approach—the Capability, Opportunity, Motivation-Behavior (COM-B) framework, used to understand behavior. The COM-B model assumes that three components interact to generate behavior: capability, opportunity, and motivation [27]. Here, capability refers to an individual’s psychological and physical capacity to engage in the activity concerned. Motivation refers to all the brain processes that energize and direct behavior (beyond goals and conscious decision making). Opportunity is defined as all the factors that lie outside of the individual that make the behavior possible. Different intervention types may change one of these three components and consequentially influence behavior.

The researchers who designed COM-B developed the behavior change wheel, which is partially utilized in the current qualitative study. The behavior change wheel presents

different behavior change components, including nine intervention functions and seven policy categories. Table 2 presents definitions of the behavior change wheel's intervention functions and policy categories upon which the current research drew its methodological and analytical choices.

2.4 Objectives

The goal of the current research is to provide an overview of barriers to the well-functioning digital market, with a focus on consumer detriment. Hence, this work's primary objective is to overview existing work about dark patterns and T&Cs and assess what measures could be applied to reduce their detrimental effects on consumers. To reach this objective, a series of sub-objectives were created, as follows.

1. To identify and re-categorize dark patterns affecting consumers' well-being
2. To understand better how dark patterns are used and what are their harmful effects on consumers
3. To identify the main barriers to understanding and taking into account information disclosures in T&Cs
4. To investigate the vulnerability concept and assess whether specific groups of consumers are more prone to manipulation by maliciously intended UI designs

Fulfilling these sub-objectives is essential because it will enable an understanding of how to ensure well-functioning digital markets. Policymakers and companies could use the current research results to establish means to decrease consumer detriment. Such means would ensure balanced information between digital companies and consumers and remove restraints and obstacles that consumers confront when transacting online.

Table 2: Definitions of intervention functions and policy categories (as presented in Michie et. al. [27]).

Interventions	
Education	Increasing knowledge and comprehension
Persuasion	Using communication to induce positive or negative feelings to stimulate action
Incentivisation	Creating an expectation of reward
Coercion	Creating an expectation of punishment or cost
Training	Imparting skills
Restriction	Using rules to reduce the opportunity to engage in the target behavior
Environmental restructuring	Changing the physical or social context
Modeling	Providing an example for people to aspire to or imitate
Enablement	Increasing means/reducing barriers to increase capability and opportunity
Policies	
Communication	Using print, electronic, telephonic, or broadcast media
Guidelines	Creating documents that recommend or mandate practice. This includes all changes to service provision
Fiscal	Using the tax system to reduce or increase the financial cost
Regulation	Establishing rules or principles of behavior or practice
Legislation	Making or changing laws
Environmental/social planning	Designing and/or controlling the physical or social environment
Service provision	Delivering a service

3 Methods

Different methods were applied to reach our research objective. First, a systematic literature review was conducted. Then, a series of expert interviews were carried out to gather qualitative insights.

3.1 Systematic Literature Review

We reviewed relevant literature about dark patterns in the context of consumer-related digital services. Only relevant literature published in English was included in the review. Books were not considered, but non-peer-reviewed publications, such as working papers and reports, were considered.

To identify suitable literature, we searched for publications from the last ten years, from 2010 to 2020. The date range started with 2010 because it was in this year that Harry Brignull, a user experience (UX) specialist with a cognitive science background, coined the term “dark patterns” [28]. To create search queries, we used the different combinations of the following terms:

misdirection, manipulation, dark patterns, terms and conditions, online shopping, consumer, review, nudge.

To accommodate the research’s multidisciplinary nature, we selected four databases: two multidisciplinary databases, namely, Web of Science (WoS)¹ and Elsevier (ScienceDirect)², and two databases that targeting more computer-related research, namely, Association for Computing and Machinery (ACM)³ and Institute of Electrical and Electronics Engineers (IEEE)⁴. The preliminary search resulted in a low number of publications. Hence, we also ran a search through the search engine Google Scholar. Additionally, when reviewing some of the publications, we applied the snowballing method to identify potential literature. That is, we found work that was not detected via our search queries but had been cited by the authors of the selected studies.

Before selecting publications for our review, we read the titles and abstracts and identified 30 relevant publications. Some of these publications did not cover dark patterns, but they were included because they would enable a greater understanding of the potential impacts that misleading designs might have on consumers and the digital market.

¹ WoS holds more than 161 million records across 254 subject areas, including the Social Sciences Citation Index (SSCI) and Conference Proceedings Citation Index (CPCI).

² ScienceDirect contains around 16 million articles from over 2500 scholarly journals. The fields of research include Physical Sciences & Engineering, Health Sciences, Life Sciences and Social Sciences & Humanities.

³ The ACM Guide to Computing Literature “includes all of the content from The ACM Full-Text Collection along with citations, and links where possible, to all other publishers in computing.” The database contains 2,853,540 bibliographic records (ACM, n.d.).

⁴ The IEEE database includes the content of “more than four-million full-text documents from subjects of electrical engineering, computer science and electronics” (IEEE, n.d.).

3.1.1 Data Extraction

Before extracting the data, we read the selected literature and divided it into three categories.

- a. Publications that directly discuss dark patterns
- b. Publications that discuss nudging
- c. Publications that might be useful to enable a greater understanding of dark patterns and issues related to T&Cs

After the first read of selected publications, we decided to extract the information about dark patterns according to a framework proposed in previous research [29] Originally, this framework was developed to categorize privacy-related dark patterns. However, because our findings extend beyond such patterns, we modified the framework to include the data fields presented below. (The data fields highlighted with an asterisk [*] are presented separately in Appendix B.)

- *Name(s)*
- *Summary* – a short description of the dark pattern
- *Context** – when a particular dark pattern might be applicable, for instance, online shopping and purchasing within a non-e-commerce application
- *Effect* – potential effects and consequences that the pattern might have on a consumer
- *Examples* – examples showing how the dark pattern might be used (Here, we present either existing examples or a hypothetical one. If possible, we use images to visualize dark patterns)
- *Related patterns or nudges** – if there are any related dark patterns
- *Categorization** – categorization described in the reviewed literature
- *Psychological effects** – description of psychological effects that could be exploited by the dark pattern

Further, we categorized the identified dark patterns into two groups: (1) dark patterns that are potential barriers to consumers' choices, or the *active choice* category, and (2) dark patterns that may decrease consumers' comprehension of the contractual content, or the *comprehension* category.

3.1.2 COM-B Mapping

After identifying dark patterns, we used the COM-B framework and mapped the identified dark patterns to each of the components that affect behavior change: capabilities, opportunities, and motivation. The mapping was performed as a group: three researchers, through discussions, mapped each of the dark patterns to one component that might be most affected by the specific pattern. However, this does not mean that a particular dark pattern will not affect the remaining components. In a similar way, we also used the COM-B framework to map the results of the expert interviews related to understanding and acting on T&Cs. However, this mapping was conducted by one researcher only.

3.2 Expert Interviews

After a systematic literature review, we collected qualitative data from experts. We used interviews to gain in-depth knowledge from multidisciplinary fields. Therefore, our interviewees were experts whose work relates to the consumer market. In particular, we interviewed experts from areas such as law, economy, philosophy, and behavioral and cognitive sciences.

Each of the eight interviewees was asked four questions. First, we asked for an assessment of which dark patterns might be most harmful to the consumer. Second, we asked about the main barriers that consumers encounter in terms of understanding and acting on T&Cs. Third, we asked about the measures that should be applied to overcome the obstacles caused by dark patterns and disclosures in T&Cs. Finally, we asked whether there were any specific groups of users who might be affected by dark patterns and contemporary presentations of T&Cs.

All the interviews were conducted online through a video conferencing platform. Each interview lasted a maximum of one hour. The interview audio was recorded, and the recordings were manually transcribed. Thematic analysis was performed on the transcribed texts. Finally, the intervention functions and policy categories of the behavior change wheel were used as the main themes to identify possible ways of overcoming the barriers (based on the responses to question three).

4 Findings

As stated earlier, this research's main objective is to review the existing knowledge about dark patterns and T&Cs and assess what measures could be used to reduce their detrimental effects on consumers. To reach this objective, it is first necessary to identify dark patterns that might be harmful to the consumer and simplify their categorization (sub-objective 1). This part of the research was enabled by COM-B mapping.

4.1 Dark Patterns

Among the 30 publications identified through the systematic literature search, only 19 directly concern digital markets and consumers. The publications that qualified for review are listed in Appendix A. In these publications, we identified 26 dark patterns that were in alignment with our research objective. Definitions and examples of these dark patterns are presented below. Some of the identified dark patterns are related to each other and exploit specific biases and heuristics. Moreover, some patterns might be more prominent in specific contexts, e.g., e-commerce mobile or web applications. Details concerning the inter-dependencies between the identified dark patterns, the context of their use, and their psychological associations are presented in Appendix B. Details of the identified dark patterns are provided below.

Activity message

Summary. This pattern informs consumers about the activity of other consumers using the same service.

Effect. It may lead to the purchase of products or services based on information that there is limited availability or that others find the product or service worth buying. If no other information is given, the consumer, misled by the presented information, may overpay.

Example. A website might contain information that many other consumers are also viewing the same product during the purchase of accommodation, but it is not necessarily true that others are viewing the same product. Under the assumption that it is true, the consumer might feel that the product is worth the price and book the accommodation as a result of social pressure, even if they initially did not intend to. Figure 1 presents an example of the *activity message* pattern on a fictional hotel website.

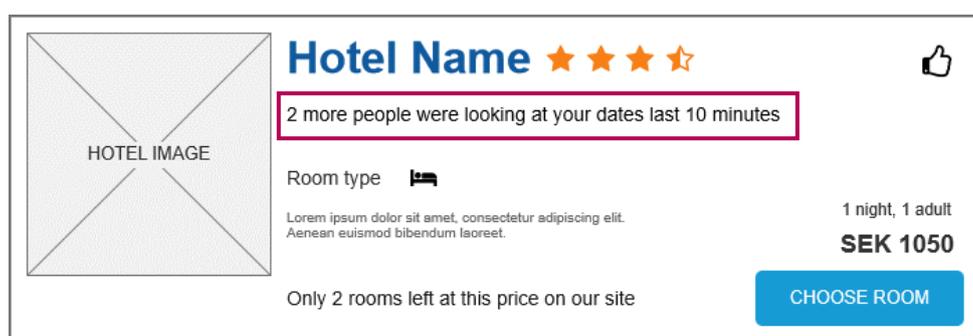


Figure 1: Fictional hotel website page showing consumers that two people considered this accommodation very recently, pressuring consumers to purchase (highlighted in red).

Address book leeching

Summary. This pattern uses contact lists that are uploaded to the service from a consumer's device. However, the consumer is unaware that their connections are being stored and processed by the service provider.

Effect. Importing contacts may result in information being exposed to different third parties and place consumers' privacy at risk.

Example. A mobile e-commerce application that has access to contacts, may share these contacts with third parties that track consumers' purchases and might spam their friends from the contact lists with suggestions for such products.

Bad defaults

Summary. This pattern applies predefined settings that users are unlikely to change. It is an inverted default nudge that was initially created to help users by pre-selection of the most beneficial options.

Effect. Various harmful effects may result from predefined options—from over-disclosure of personal information to pressuring consumers to purchase particular products or services. Pre-selection may result in economic loss; for instance, the consumer might not be aware of charges for an ongoing subscription after the trial period.

Example. Pre-selected or more visible options related to the product pressure consumers to buy more expensive, associated products (see Figure 2). This dark pattern also concerns subscriptions, for instance, when a consumer signs up for a new service and the option for an automatic renewal is pre-selected.

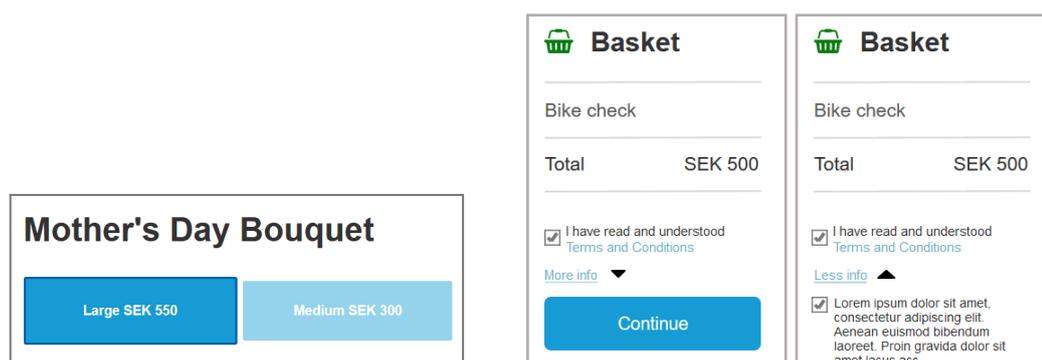


Figure 2: On the left: the default pre-selected price is higher than the alternative; on the right: the pre-selected option is hidden.

Bait and switch

Summary. In this pattern, the customer is shown that a specific action will lead to a particular result, but the decisional outcome is different from what was expected.

Effect. This pattern may result in an unexpected decisional outcome that is brought about by switching consumers' attention to purchasing a product different from the one initially desired. This may potentially result in economic loss.

Example. The consumer might want to purchase a good or service that was advertised. However, during the purchasing process, they are exposed to a series of ads related to purchases that the consumer did not intend to buy. For example, the ticket purchase presented in Figure 3 shows that a ticket was ordered for an exhibition, but a set of razor blades at a discounted price was offered to the consumer as an optional purchase.

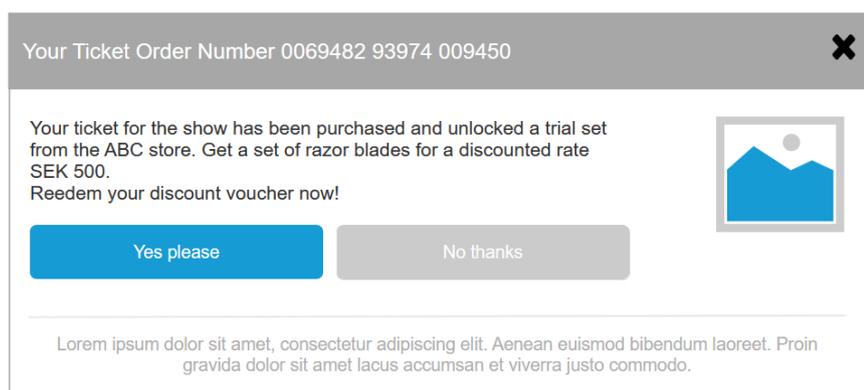


Figure 3: Example of the Bait and switch pattern in which a consumer is presented with an optional purchase while purchasing another service/good.

Confirmshaming

Summary. This pattern makes the consumer feel guilty about not opting into something or opting out of something. It uses the power of language to steer consumers into making a specific and undesired choice.

Effect. It plays with consumers' feelings by making them feel dishonorable or stupid.

Example. Confirmshaming might be used when a customer makes a purchase and they are offered a discount. In this pattern, the language used to present the option makes consumers feel guilty about not opting for the deal, as shown in Figure 4.

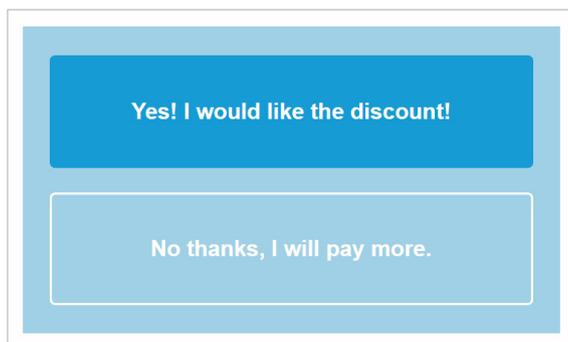


Figure 4: Example of Confirmshaming where a consumer is asked whether they would like to opt for the discount.

Disguised ads

Summary. Ads that consumers may see but do not realize are ads.

Effect. Consumers might click/tap on such hidden advertisements, and as a result, they may be exposed to unwanted ads and motivated to purchase additional products. Consequently, they might experience economic loss.

Example. On a website distributing software downloads, ads might be disguised in the UI design so that they appear identical to actual software download buttons. Once the user clicks on the download option, instead of the software being downloaded, users are redirected to websites related to the ad.

False hierarchy

Summary. This pattern makes some of the options appear more prominent than others. It is also known as *pressured selling*.

Effect. This pattern can prime the consumer to select initially undesirable settings, and this negatively impacts privacy and places consumers at potential risk. It may also lead to financial losses.

Example. In an e-commerce application containing recommended settings, such as delivery options, consumers might be prone to believe that the most expensive delivery is the best option, as it has been most prominently recommended.

False urgency

Summary. This pattern indicates to consumers that only limited quantities of products are available, and it is also referred to as *low stock*. In some situations, this is not treated as a dark pattern but as a category of dark patterns.

Effect. Limiting the amount of time spent on the purchasing decision may result in less informed choices. Consumers may buy products only because these might soon be unavailable, instead of products that fulfill their needs or match their preferences.

Example. This pattern might be used on any website presenting consumers with a choice in which products are limited in stock. However, this information is false, and it is delivered to consumers only to nudge them toward a purchase (Figure 5).

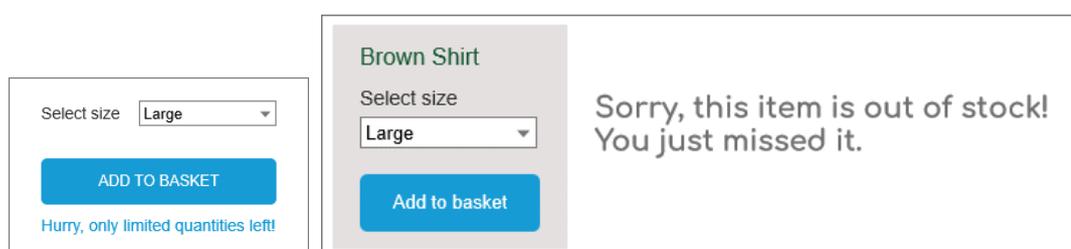


Figure 5: Examples of low-stock messages based on samples presented in Mathur et al. [30].

Forced action and timing

Summary. This pattern forces consumers to make choices on the spot. In some cases, it is considered as a category of dark patterns rather than an individual dark pattern.

Effect. This pattern nudges consumers to agree to all the T&Cs when purchasing a product or service. As a result, consumers may blindly accept all the terms and be unaware of potential risks or additional choices they could have made during the purchasing process.

Example. During the installation of a new, paid-for mobile service, consumers might be outside of their “comfort zone.” They might be traveling, limited via external factors (e.g., the discomfort of being in a public space), or under time pressure (e.g., when they perform the task in transit). In such situations, if information about the T&Cs is presented as a long, scrollable text, consumers are unlikely to read any of the information. Instead, they prefer to accept the given conditions blindly.

Forced continuity

Summary. This pattern is present when consumers are coerced to continue an action in relation to a web product. It is also known as the *hidden subscription*.

Effect. The customer is faced with unexpected financial loss.

Example. This is an example of a website that hides subscription information. Figure 6 shows the purchasing process for a fictional online wine retailer. One of the offers is the addition of a subscription if free shipping is selected. However, the details of the offer and the cost are hidden and only accessible on clicking on the inconspicuous link “Learn more.”

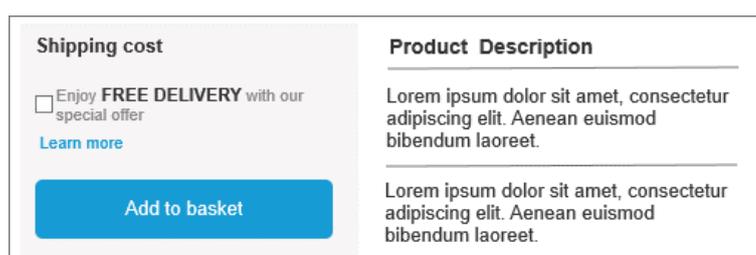


Figure 6: Example of Forced continuity through hidden subscription. Based on the findings from Mathur et al. [30].

Forced enrollment

Summary. This dark pattern coerces consumers to create accounts or share their information to complete their tasks and is also known as *forced registration*.

Effect. By forcing consumers into registration, companies may collect extensive data about consumer behavior. Because account creation is not a primary task, it is unlikely that a consumer will read the details of the account creation agreement or change any settings. As a result, consumers’ privacy might be at risk.

Example. Some online services exploit the account creation process and oblige consumers to receive emails with special offers. They may also force consumers who only visit the website to create an account, even if they do not purchase anything, as illustrated in Figure 7.

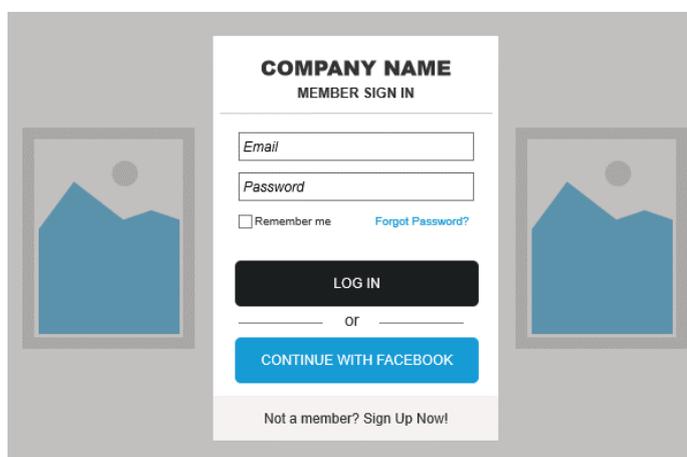


Figure 7: Example of Forced enrollment: the user cannot access a website (grayed-out) unless they log in or sign up.

Friends spam

Summary. This dark pattern deceives consumers into spamming their friends to join a service. The consumer is, however, not aware that their information is being used to spam their friends.

Effect. Spamming contacts with unwanted messages might be annoying. Additionally, it has the potential for privacy violation because the friends did not consent to share their data.

Example. In one of the well-known cases, LinkedIn asked for consumers' contact details that were afterward utilized to spam people from the contact lists.

Hard to cancel/roach motel

Summary. This dark pattern makes it almost impossible to get out of a situation. It is also known as *hard opt-out*.

Effect. Consumers have difficulty with canceling an account; thus, they give up and accept that the account exists. They might be unaware that their information is being used for undesired purposes.

Example. In some cases, the account cancellation process might become cumbersome. For instance, some companies require consumers to contact customer support and call a dedicated department to cancel an account or order. Using a different medium (e.g., phone) to contact the right services is an obstacle, and consumers might be repelled by the perceived additional effort that is required (Figure 8).

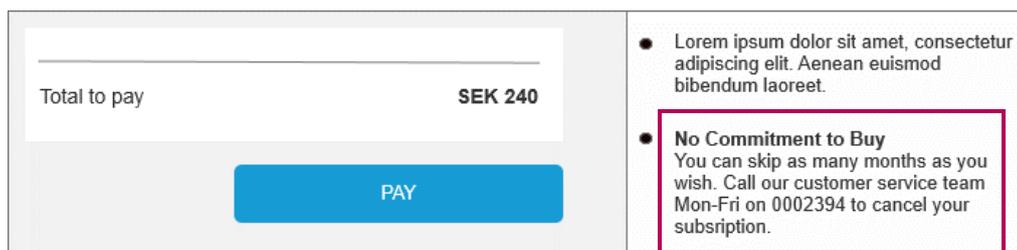


Figure 8: Example of Hard to cancel/roach motel dark pattern (According to the cancellation details presented in the side panel [highlighted in red], users are required to call the company during a specific time.)

Hidden cost

Summary. In this pattern, consumers find out about additional costs when they are already deeply engaged in the buying process.

Effect. If the hidden cost is unnoticed, consumers may pay for a service or product that they did not initially desire to purchase.

Example. Figure 9 presents an example of this dark pattern. The flower shop website presented allows consumers to select the desired product, but at the check-out, unexpected charges are added to the shopping cart, such as care and handling charges.

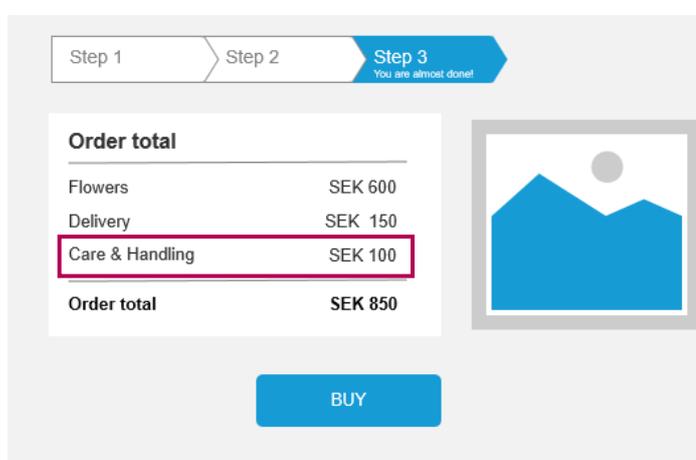


Figure 9: Example of the Hidden cost dark pattern wherein consumers are presented with an additional cost at the last step of purchasing (highlighted in red).

Hidden legalese stipulations

Summary. The service's T&Cs or privacy policies are presented to the consumer in an unfriendly way and usually mask some information.

Effect. The consumer may be unaware of the potential implications of agreeing to T&Cs, which might not necessarily have the consumer's best interests in mind.

Example. One of the T&C clauses of a gaming company stated that they would own their customers' "immortal souls." Because the terms and conditions were inappropriately displayed, the company ended up owning thousands of "souls."

Immortal accounts

Summary. Consumers are unable to delete previously created accounts.

Effect. The UI might be designed in a manner that prevents the consumer from deleting their account. As a result, the personal details provided to create an account may stay forever with the service provider.

Example. Some services do not allow consumers to delete their accounts. Instead, they only allow for deactivation of accounts. For instance, a known company that offers reference management software does not allow users to remove personal information from their database altogether. As a result, it is impossible to open a new account with the same email address.

Intermediate currency/pseudo currency

Summary. This pattern makes consumers unaware of how much money they have spent by introducing pseudo-currency.

Effect. Consumers may spend money even if they initially did not plan to spend any.

Example. Consumers might buy pseudo-currency when playing a game to spend it on game-related upgrades and similar purchases. Because the new currency does not contain explicit information about how much of the real currency it is equivalent to, consumers may overspend.

Price comparison prevention

Summary. Marketers use this dark pattern to prevent customers from making direct comparisons between products or services.

Effect. This pattern prevents the consumer from choosing the option that might be the most optimal in terms of price. Thus, the consumer may pay a higher price for a product that could have been obtained at a lower price elsewhere. Such designs based on the prevention of price comparison have adverse effects on the free market concept.

Example. This dark pattern may be used by companies when they display a product description and price that cannot be copied. Instead, the consumer is required to memorize both the product's price and its name and retype it in the search engine or dedicated price comparison software. This pattern seems to be common in mobile commerce applications. Figure 10 presents an example of an android e-commerce application that prevents consumers from copying item descriptions, thus forcing them to stay in the app and shop in their store.

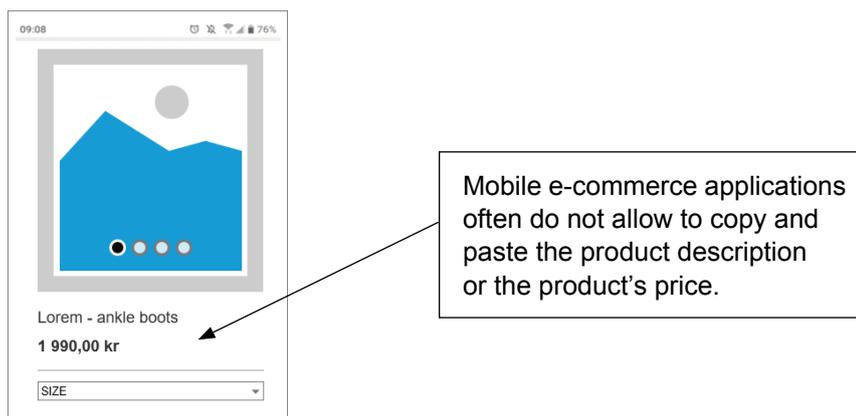


Figure 10: Example of Price comparison prevention.

Privacy Zuckering

Summary. This dark pattern is prevalent in services that collect more data than users agreed to and, also, share such information with other parties.

Effect. The way information is presented, the terminologies used, and the user experience will prevent the consumer from making any changes. As a result, the consumer may lose control over their personal information and place themselves at risk by not understanding how such information is being used.

Example. One of the most prominent examples comes from the non-commerce website - Facebook. The old version of Facebook privacy policy contained overly complicated privacy settings that prevented users from appropriately selecting their preferred settings and, consequently, they would disclose their personal information without intending to.

Rewards and punishment

Summary. This pattern incentivizes consumers into decisions that they did not intend to make but are in keeping with desires of the company. Often, rewards are presented as an additional functionality offered by the service, while punishment is usually related to “take it or leave it” situations.

Effect. Consumers are tricked into choosing something that they did not intend to select.

Example. Companies may confront consumers with request-like messages when they are trying to opt out of services. Figure 11 presents an example from a service that offers personalized ads. The consumer is forced to accept tracking in order to make ads personalized, or they will be unable to have any control over ads.

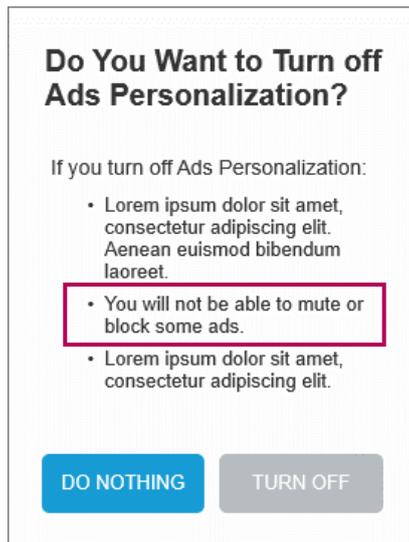


Figure 11: Example of the Rewards and punishment dark pattern (highlighted in red).

Sneak into basket

Summary. This dark pattern adds items to a consumer's shopping carts.

Effect. Companies benefit from tricking consumers and relying on consumers not having noticed the additional items added to their cart during the purchasing process. This might result in consumers facing a financial loss.

Example. One of the examples presented by Mathur et al. [30] shows a flower shop website. After the desired product is selected, a greeting card is added to the basket (Figure 12).

Your Basket		
Item	Qty	Price
 Plant Lorem ipsum dolor sit amet, consectetur <small>Lorem ipsum dolor sit amet, consectetur</small>	1	1200,00 kr
 Greetings Card Lorem ipsum	1	150,00 kr

Figure 12: Example of Sneak into the basket pattern on a fictional flower shop website wherein a greeting card is added to the basket of a customer who has purchased a plant.

Social pyramid

Summary. This dark pattern obligates consumers to recruit others to start using a service. It is also classified under friend spam and address book leeching.

Effect. Consumers are forced to share information about others, and this places their data integrity and privacy at risk.

Example. In the gaming industry, consumers might be incentivized to request their friends to join the platform. For example, mobile games may incentivize users to invite their friends to the game by making some features or goals inaccessible without online friends who also play the game (Figure 13).

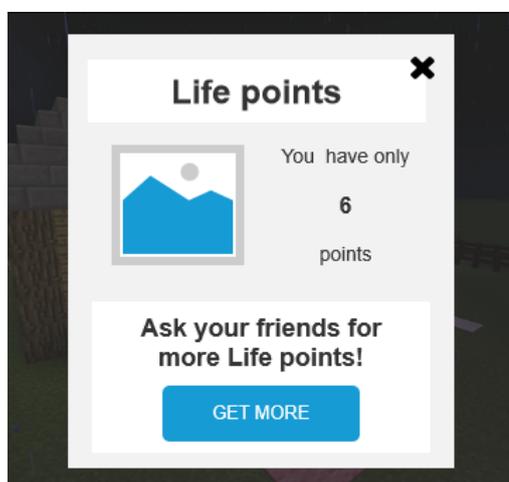


Figure 13: Example of a mobile game that nudges players to spam their friends.

Testimonials

Summary. Overall, testimonials are not a dark pattern. However, some services present consumers with testimonials of unknown origin, and such fictitious testimonials become a dark pattern.

Effect. Positive yet fictitious testimonials may nudge consumers to purchase a product or service that is defective.

Example. Mathur *et al.* [30] identified the same testimonials that appeared for different products on different e-commerce websites under different reviewer names. This indicates that these testimonials and reviewers are fictitious.

Time countdown

Summary. In this pattern, the user is informed that a special deal or discount on the product will expire soon. However, in reality, when the expiration time approaches, the same offer is still available. This is also known as a *limited time message*.

Effect. This pattern may influence the consumer's emotional state and lead them to believe that they will lose the opportunity to purchase the product. As a result, consumers may experience economic loss from making unwanted or not-needed purchases.

Example. Mathur *et al.* [30] identified many examples of time-limited offers presented to consumers. However, in all instances, these offers remained on the websites after the time count ended. Either the request was restarted when the consumer refreshed the site or the offer re-appeared on the page daily. There were also instances where the time-limited offer was presented without a finite time of when it would expire. Figure 14 illustrates examples of the timer countdown.

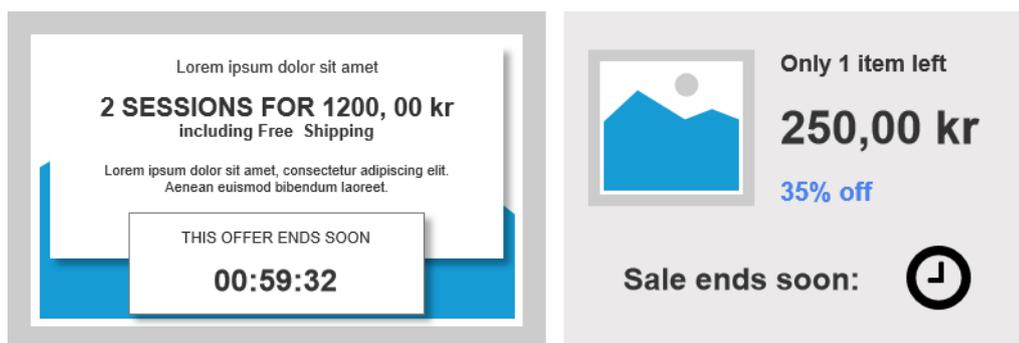


Figure 14: Examples of the Time countdown dark pattern used in different formats.

Toying with emotions

Summary. Language, style, color, and other UI design elements can be used to evoke a particular emotional state. These design elements are applied in different ways to persuade the consumer into specific action. This dark pattern relates to Confirmshaming.

Effect. This pattern may result in unexpected purchases or limit the consumer's autonomy by tricking them into not changing the current settings.

Example. Companies may offer different choices for purchasing a product, as shown in the example of the fictional book store illustrated in Figure 15. The option for not purchasing the Easy reader version is framed negatively. Therefore, the consumer is more likely to choose Easy reader because it is presented as a more beneficial option. However, the consumer might actually prefer to have a print version of the book.



Figure 15: Example of the Toying with emotions dark pattern in which the option preferred by the company is presented as the more beneficial one.

Trick questions

Summary. This pattern is usually formed as a question that appears to be one thing while meaning something else. This dark pattern may rely on confusing wording, double negatives, or other similar tricks that could confuse consumers.

Effect. Consumers may unwittingly agree to different kinds of subscriptions, communications with service providers, and similar options.

Example. A fictional company called ABC (Figure 16) uses confusing terminology and choice options to trick consumers during the signup process (during which consumers might opt in for or opt out of special offers).

ABC

Sign Up Form

First name

Second name

Email

Password

Please do not send me details about products or special offers from ABC.

Please send me details about products or special offers from third parties recommended by ABC.

SIGN UP

Figure 16: Example of the dark pattern Trick questions (highlighted in red).

4.1.1 COM-B Categorization

The first sub-objective of this research was to identify and categorize dark patterns that affect consumers. Past research has revealed multiple categorizations of dark patterns, and some of them are applicable to the dark patterns identified in our research: social proof, sneaking, urgency, misdirection, scarcity, obstruction, forced action, nagging, interface interference, maximize, publish, obscure, and deny (see Appendix C for definitions). Additionally, dark patterns can be categorized according to their strategic function and harmfulness. The categories according to strategic function are sales, data gathering, views, time spent on the product/service, and miscellaneous, and the categories according to harmfulness are just annoying, moderately bad, and need official regulation. Such a variety of categories makes it challenging to map them onto a behavioral change wheel. Hence, we aimed to simplify the categorization of dark patterns by sorting them according to consumers' tangible and cognitive activities. Table 3 presents the names of the identified patterns and our categorization results according to their potential effects on *active choice* or *information comprehension*.

According to our mapping, most of the dark patterns affecting active choice tend to influence motivation (11); only three influence capability; and only two affect opportunity. However, within the category of dark patterns that are more likely to influence comprehension, the mapping appears less unidirectional. Here, four categories

influence motivation; six influence capability; and three, opportunities. Note that apart from the associations mentioned here, both individual dark patterns and the two newly identified categories may affect the COM-B framework's multiple components. We based our mapping on whether a given category/dark pattern may have the strongest influence on capability, opportunity, or motivation.

Table 3: Dark patterns mapped onto the COM-B framework.

	Capability	Opportunity	Motivation
Active choice	Bait and switch Intermediate currency Sneak into basket	Forced continuity Price comparison prevention	Activity message Bad defaults Confirmshaming False hierarchy False urgency Hidden cost Rewards and punishment Social pyramid Testimonials Time countdown Toying with emotions
Comprehension	Hard to cancel/Roach motel Address book leeching Friends spam Hidden legalese stipulations Privacy Zuckering Trick questions Disguised ads	Forced continuity Forced enrolment Immortal accounts	Bad defaults Social pyramid False action and timing Toying with emotions

4.1.2 Interview Results

To better understand how companies use dark patterns and how harmful these tricks might be for consumers, which is the second sub-objective of this research, we conducted a series of expert interviews. The responses from the interviews were assessed to draw out a few distinct themes based on the harmfulness of dark patterns, as presented below.

“Second-generation” dark patterns

Our experts were most concerned about dark patterns that are difficult to describe and hard to identify by consumers. The “traditional” dark patterns, or what one of our participants referred to as *first-generation* dark patterns, are easier to recognize by consumers and, consequently, might be easier to regulate. Companies applying first-generation tricks are aware that they might be annoying for the consumer and are, therefore, unlikely to overuse them. As explained by one of our experts, “It makes you accept something you are not willing to accept. I think that it is a problem, but it is more a problem that has a tendency to balance itself because people will neglect or decline from revisiting shops that use these too aggressively” (P4). On the other hand, *second-generation* dark patterns are more dangerous because the effects of this category of

dark patterns are less observable and prevent self-regulation of the market. These dark patterns are represented by tricks that are insidious and “harder to get rid of because they are slippery, and hard to define” (P3), for instance, the usage of specific wording that is deceptive but seems honest and is hard to argue against: “[...] regulating that sort of thing is much harder because knowing how to write it, in a way that is specific enough and open enough is hard” (P3).

An example of a second-generation dark pattern can be derived from the hotel industry, where a hotelier might have different types of rooms and applies one of the urgency patterns. That is, they encourage consumers to buy because there are only a limited number of rooms of a certain type left. However, the hotelier might have a range of different room types that are almost identical but have been segmented to make it seem like there is always low availability of a particular type of room. Simultaneously, a specific algorithm might be applied to ensure that these room types are always prioritized and shown to the consumer first. Therefore, considering all the manipulative tricks embedded in such an approach—similar room types, false urgency, algorithmic manipulations—it becomes almost impossible for consumers to understand what is happening and challenging for regulators to identify.

Exploiting data. Our interviewees also mentioned that the dark patterns which cause the most damage to consumers are those that combine consumer’s information, such as behavioral data, various types of personal information, and traditional dark patterns. Once again, they are difficult to identify, and this is what makes them more harmful than others. “[T]he company has a lot of data about you, so they know about you as a person, plus they know about human behavior in general. And you cannot see it, you cannot see through it [design]” (P1). These dark patterns result in an unequal relationship between consumers and businesses, in a way that companies can truly take advantage of consumers by exploiting their vulnerabilities.

Useful vs. useless information. Some dark patterns are hard to define as harmful because some of the information provided might be useful to the consumer. For instance, “it can be useful to know that there are not many places left when you are booking somewhere, and you should make your mind up” (P4). However, such a design might be malicious, as it might only prevent consumers from shopping around, thereby negatively affecting the probability of the consumer making the most optimal choice.

Targeting vulnerabilities. Some of our experts talked about second-generation dark patterns that target specific vulnerabilities. This pattern is closely related to the concept of personalization. It is difficult to judge the act of collecting personal information and using it to customize services as good or bad. This poses a challenge for consumer agencies, as one of our experts mentioned: “some consumers I think, they see it as a good service that a company knows when I have been looking for a certain product, and it will be at a lower price, and then I suddenly get that offer on my computer” (P5). However, it must be taken into account that extensive data collection might feed on consumer weaknesses and target them. Companies may possess information about consumers’ past behaviors, personal information related to their economic status, or more in-depth information: “if you are targeting those addicted to gambling, for

example, by using those big buttons containing slogans like ‘try your luck’” (P1). The targeted consumers may be bombarded with customized offers of products and services that, at the time, seem to be valuable and worth having. Yet, in reality, consumers do not need them.

Economic loss

Our interviewees agreed that among the dark patterns, the most harmful ones are those which, in some ways, result in consumers’ economic loss. These dark patterns mostly prey on behavioral biases and often carry hidden or unexpected costs, e.g., “you have to read the fine print so that you turn up with your printed boarding pass; otherwise, they will charge you some fee that has nothing to do with cost recovery” (P4).

With regard to dark patterns that cause economic loss, our participants mostly discussed traditional, first-generation dark patterns, such as scarcity-based patterns, drip pricing, price partitioning, subscription fees, hidden costs, and so on. These dark patterns might be less challenging to identify and easier to regulate, but it does not mean that they are ineffective or harmless. In addition, they might still be challenging to spot, as some of these tricks are visible only after consumers create accounts with companies and are not visible to an enforcer scrolling through the website. Thus, despite consumer protection regulation, without appropriate methods to surveil companies, it is hard to assess whether the company is violating any rules by manipulating consumers’ choices.

Privacy violation

Some of the interviewed experts mentioned dark patterns that violate privacy. These dark patterns are closely related to second-generation patterns and exploit extensive data collection. They are also tied to the inefficient disclosure of information in T&Cs. People are often unaware of data collection and processing practices because they might lack information about it or might not understand how the technology works. “I think people do not understand the implications of it. They say, ‘I got nothing to hide,’ ‘I don’t care seeing adverts at all.’ They don’t understand the idea of the profile being built on them individually and how it might be misused” (P3). One of the problems related to extensive data collection is the fact that digital markets are based on the financial framework. Such a framework involves a contract between two parties and space for remedies: if something goes wrong, the consumer can request compensation. Yet, it is impossible to give the data back, as it has been already processed and its benefit is rarely quantifiable.

Restraints

Some of the other more harmful dark patterns mentioned are restraining patterns. The following restraining patterns emerged from the interviews:

Locking in one service. Some dark patterns aim to lock consumers into one service. These are dark patterns that make it hard to cancel the service or make it difficult to unsubscribe, as they simply “get you to lock in the purchase and not let you shop around” (P3).

Hiding choice. Our experts perceived dark patterns that deter consumer choice as harmful. For instance, in the context of personal advertising, emotional language is sometimes used to deceive consumers: “when you are trying to opt-out from providing data for personalized advertising, you see ‘You will get less relevant results.’ While what it exactly says is we will not use your data to target you or segment you based on different things, or try to predict what you might want to have” (P7). Other examples that the experts provided were related to bad defaults, where predefined options restrict choice by preying on the status quo effect.

4.2 Terms & Conditions

The third sub-objective of this research was to identify barriers to understanding and acting on T&Cs. Hence, below, we first discuss some of the past research about T&Cs and then present insights from the expert interviews.

T&Cs encompass an agreement that defines the relationship between the user and the online service provider [31]. One of the potentially unfair clauses in consumer contracts is “having a consumer accept the agreement simply by using the service, not only without reading it but even without having to click on ‘I agree/I accept’” [32]. To the best of our knowledge, the topic of T&Cs in the context of the digital market has not been widely researched. However, considering the unfair clauses mentioned above, it seems reasonable to assume that the barriers to understanding T&Cs are similar to the obstacles concerning the comprehension of privacy policies and End User License Agreements (EULAs).

Past research shows that people do not read privacy policies and, instead, blindly agree to the terms of online service providers. However, some past studies showed that when people are, by default, confronted with policies, they may devote longer time and greater effort towards reading lengthy texts [31]. Regardless, the biggest obstacle to reading and understanding privacy policies, EULAs and, presumably, T&Cs is the length of such documents and, often, their difficult wording and legalese. Such documents frequently consist of statements that go beyond the functional literacy threshold, as shown in the research on energy suppliers’ terms in the UK [33]. Past research reveals that simple changes in the visual representation of such text may improve users’ comprehension. For instance, paraphrased EULAs (presented as bullet-point statements written in everyday language) result in more positive attitudes among users and increase the exposure time compared to traditional EULAs (presented as a long legalese text) [34]. Additionally, abbreviated information across multiple displays that users are confronted with has comparable effects.

Companies frequently utilize T&Cs as a means to convey important information. However, this approach is not recommended by international organizations, for instance, by the OECD. In their report [2], the OECD recommended that appropriate online disclosures (here, it refers to information disclosed by a company and required by a consumer to make an informed choice) should meet the following criteria [2]:

- information should be simple, straightforward, accurate, and relevant;
- businesses should not rely on T&Cs to convey critical information;

- companies should obtain express and meaningful consumer consent;
- images, audio, and video, and not just text, should be used where appropriate;
- *when* consumers receive information is of crucial importance;
- personalization could potentially be used to improve disclosures;
- complex information should be presented in a way that enables analysis by technology and third parties.

Some empirical studies showed that the default exposure to the text of T&Cs and presenting it in a simplified and shortened manner, may increase the readership and comprehension of the terms [35]. Further, some visual cues, such as cost cues (informing the consumer about the amount of time required to read T&Cs), seem to encourage consumers to read agreements [35]. Similarly, an experiment carried out by the Danish Competition and Consumer Authority [36] showed that, indeed, simple changes in the display of information related to T&Cs influence consumers' attention and comprehension and could lead to behavioral change. For example, research shows that T&Cs presented up-front (instead of T&Cs hidden in secondary layers of UI design, e.g., hyperlinked text that opens on a new web page) and in a simplified manner (with short but precise information, relative to consumers' activity) result in consumers noticing T&Cs and acting upon them. Moreover, such designs increase the ease with which consumers are able to identify specific terms and decrease the time needed to navigate through the T&Cs interface. Additionally, consumers presented with simplified T&Cs seemed to be more satisfied and expressed lesser frustration than consumers presented with standard formats of T&Cs.

Similar research has been carried out in the UK, where researchers measured understanding of and engagement (opening of) with T&Cs and/or privacy policies [26]. Based on multiple experiments, the research produced a set of guidelines and best practices that could be applied to increase information comprehension when presenting consumers with the terms of an agreement. Some of the most successful presentation guidelines are shortening information to display only the key terms, using a scrollable text box instead of a clickable link, using timely and short texts instead of one lengthy document, and using icons and other types of illustrations to support text [26]. Other additional but less effective ways of presenting information have also been identified. For example, methods to increase engagement with T&Cs, particularly methods to increase the opening of privacy policies, were found to be less effective. The UK research has shown that providing consumers with cues related to the length of time required to read the privacy policy, as well as urging consumers to read by indicating that it might be their last chance to do so, resulted in an increase in the rates of opening of policies. However, the UK research did not investigate ways to improve engagement with T&Cs.

Dark patterns influencing T&Cs. Some of the dark patterns relate to T&Cs and may have a detrimental effect on consumers, for instance, dark patterns that decrease understanding of the relationship between the consumer and service provider. In this research, we found that *Hidden legalese stipulations*, *Privacy Zuckering*, *Hard to cancel*, *Bad defaults*, *Forced enrollment*, *Immortal accounts*, and *Forced continuity* are dark patterns that might directly affect the understanding of T&Cs, e.g., the rights that consumers have concerning interaction with a given service. The recommendations of

the OECD with regard to how information should be disclosed to consumers may help overcome barriers that the above-mentioned dark patterns cause.

4.2.1 Interviews Results

A few main issues surfaced in the interviews responses regarding the most significant barriers that consumers encounter when interacting with T&Cs. The expert's answers were relatively uniform, resulting in the identification of the following barriers:

A few main issues surfaced in the interview responses with regard to the most significant barriers that consumers encounter when interacting with T&Cs. The expert's answers were relatively uniform, and the following barriers were identified:

- *Length.* Currently, T&Cs are too long and, therefore, “disrupt the pace of the transaction” (P2). The lengthy documents require consumers to dedicate a large amount of time to gain a full understanding of the T&Cs. However, consumers are usually time-poor in the digital market context, due to external constraints, such as the semantic context, or because reading and understanding T&Cs is not their primary task. Their primary task is to complete the transaction, and not to spend time seeking information in lengthy texts, which contain information that might not even be useful.
- *Information overload.* From an economic perspective, decisions based on a larger amount of information should result in a more optimal decisional outcome. However, this is not the case when people are constrained by multiple factors, as in the case of interacting with technologies, especially in the context of digital markets. Therefore, there is a need to make T&Cs more comprehensible and provide consumers with less but more useful information. As P8 states, “not having the information in a kind of succinct, snappy, short headline summary is probably the number one barrier to engagement.”
 - *Lack of prioritization.* It is important to provide only essential information and prioritize it: “a lot of thinking needs to go into that because you could obviously highlight flattering information, and you could relegate and bury less good information for the consumer.” Therefore, lack of prioritization is also an important issue. In this regard, our respondents stated that traders might not know what information is more critical. Here, enforcers could help by identifying priority information. However, the enforcers' recommendation may not always match consumers' needs. For instance, when it is based on consumer' complaints gathered by consumer agencies, “it's not always that complaints match what we as a consumer protection agency see as major problems or issues.” On the other hand, such complaints may not entirely reflect the actual complaints that consumers might have because they are more likely to be addressed directly to service providers than to consumer agencies.

One of the solutions about what information should be disclosed, as suggested by our experts, was acquired from Sweden's insurance law. Here, precise requirements provide consumers with a short and comprehensible label containing all the needed information about a product. “I think it would be possible to do the same for selling any product. I think you have to know what you

buy, and what it costs, and if there are some things that this product does not have that other products usually have” (P6).

- *Understandability.* Some of the respondents mentioned that T&Cs are still, for some reason, written by lawyers, unlike in other fields, where intermediary experts are hired to design products or services for people in order to make them functional and understandable. “[...] people who train in law lose sight of what people that do not understand the law as an academic discipline, how they understand information. So, that is perfectly fine; it happens all the time. [...] That is why engineers don’t design buildings, architects design buildings. But for some reason, that has not filtered down into law. For some reason, we allow lawyers to design law, instead of just informing law” (P5).
- *Difficult language.* The language often used in T&Cs is complicated and filled with jargon, and this makes the terms difficult to comprehend. There is a need for “the simplicity of the language, and just writing things in very plain terms, without any jargon, without any complexity, and actually not writing paragraphs but just really specific sentences or points” (P8).
- *High reading score.* Another issue related to understandability and language used in T&Cs is that some of the current T&Cs seem to be too difficult to read by average users, “[...] we actually looked at all of the search engines’ T&Cs [...] what is the reading age in English roughly, and what is the complexity broadly. And the vast majority of platforms came out as very complex and having a very high reading age” (P7).
- *Unusable.* Currently, T&Cs have no real meaning for the consumer. That is, they are a secondary source of information, and consumers have no reason to pay attention to them: “in a market where things often go wrong, people would pay more attention to T&Cs” (P5). Further, “no one is able to use them, and very few people interact with them” (P5). To improve the interaction with T&Cs, they must become more relevant to the consumer.
 - *Lack of choice.* T&Cs are a contract that is not negotiable: “in a lot of markets, it is legitimate consent that if you read it, I mean, you say yes or no. So, if you want to do something, if you want to use the service, you don’t have much choice” (P4). This calls for providing consumers with more options that, in turn, could potentially impact their engagement with T&Cs. As P3 stated, “Should the consumer be forced to accept this massive monolithic document with loads of information in it. Should it be broken down, should there be ways of not accepting bits of the contract.”
 - *Lack of meaning.* T&Cs might not be worth engaging with because they hold no meaning for the consumer. If they had a material effect, for instance, if the consumer was allowed to control whether they want to have a return policy for 60 days, 30 days, or no return policy, T&Cs might gain more meaning. “If there is a reason to engage in the actual decisions that you make in relation to T&Cs that have a material effect, then absolutely. I think that is a really good idea” (P8).
- *Visual representation.* Most of our experts considered the visual design of the current T&Cs as problematic.

- *Legal defensiveness.* The use of the tick box to acknowledge T&Cs is “the biggest lie on the internet” (P3). People do not look at the terms and do not even open them. In some instances, e.g., when buying new software, consumers are forced to scroll through the terms and then acknowledge them. “[I]t is probably a form of legal defensiveness, so they can say ‘look, we have done everything we can, we forced the user to look at it, there is nothing else we can do’” (P3).
- *Lack of clearly defined patterns.* There is no information available to companies regarding the best practices for informing consumers through visual design. Yet, empirical evidence shows that presenting information as bullet points, questions and answers, etc., increases comprehension and engagement.
- *Secondary importance.* T&Cs are often secondary to the interaction. There are other, more prominent visual cues that people attend to when purchasing online, instead of engaging with T&Cs. “[P]eople take [...] fancy brand, clear fonts, a nice simple, clean page structures, all of these kinds of layout factors, potentially fake reviews, potentially pressure signals; they take all of these kinds of things into their purchase decisions probably much more so than they take into account the delivery, return cost, or the privacy or other kind of aspect of T&Cs they are maybe signing up to too” (P8).
- *Lack of actionability.* The current designs of T&Cs, in most cases, do not enable actionability beyond a binary choice. Depending on the context and variability in the terms, actionability could improve consumers’ decisions. “If there is a lot of variation, then I think we can make people care about it if we present it to them in a way that is very actionable, makes it very easy for them to understand the differences between different sellers, in terms of how good the T&Cs are for that specific product” (P5).
- *Too many contracts.* Consumers are faced with far too many contracts in the digital market, and this makes it difficult for enforcement agencies to check whether these contracts are lawful and fair. “I think even if you banned a lot of terms and conditions, the enforcement is still going to remain an issue” (P4).
 - *Habituation.* A number of contracts that consumers have, together with the current most popular visual representations (i.e., ticking a check box), have instilled a habitual behavior among consumers. This behavior is one of the barriers to consumer engagement with T&Cs. “[W]e have already trained consumers to avoid or not interact with them [...] I think you would have to sort of radically re-invent how they work and how they function, how they are presented to consumers if you want them to have any role in guiding consumer behavior” (P5).
- *Lack of empirical evidence of the effects of T&Cs on behavior.* It seems that most of the research on T&Cs is related to measurements of understanding or engagement as outcome variables. As pointed out by P8, we assume that understanding and engagement will reflect in behavior, but “we do not know a huge amount about it.” This calls for further research to assess whether people will consider T&Cs if they are presented differently from the current meaningless acknowledgments and whether they will act upon such considerations.

Overall, the above list presents the main barriers to understanding and acting on T&Cs, according to the eight experts who were interviewed. From their responses, it

seems that, in general, information disclosure through T&Cs is ineffective and should be considered as a *policy failure*, as described by one of our experts: “[...] everyone is just extremely busy and time-pressed, and it is a completely ineffective way of communicating with consumers, and I think that businesses use that” (P4).

4.2.2 COM-B Categorization

We categorize the barriers listed in the previous subsection according to the COM-B framework. Table 4 presents the results of our categorization. Similar to the mapping of dark patterns, a specific barrier is categorized according to whether it may affect capability, opportunity, or motivation. However, to make the categorization more useful, we tried to determine which of the COM-B components are affected the most by each barrier.

Table 4: Barriers to understanding and acting on T&Cs mapped into the COM-B framework (sub-factors are italicized).

Capability	Opportunity	Motivation
Length	<i>Lack of prioritization</i>	Unusable
Information overload	<i>Lack of choice</i>	Meaningfulness
Understandability (<i>Language, high reading score</i>)	<i>Lack of actionability</i>	Secondary importance
Visual representation	<i>Legal defensiveness</i>	Habituation
Too many contracts		

4.3 Consumer Vulnerabilities

The last sub-objective of this research was to investigate the vulnerability concept in order to assess whether specific groups of consumers are more prone to manipulation by maliciously intended UI designs. Hence, we want to understand better who is most vulnerable to dark patterns and inefficient representations of T&Cs. Following this, we focus on the findings related to our main objective—identifying measures that could improve consumers’ informed choices and, thereby, ensure a well-functioning digital market.

In our interviews with the experts, we asked them to identify potentially vulnerable groups that are affected by the dark patterns and barriers identified in the previous sections. Overall, our interviewees admitted that they had little practical experience with vulnerable consumers. However, some of the experts mentioned the following potentially vulnerable groups:

- *People with cognitive impairments.* These individuals might be more prone to the effects of dark patterns. “[I]f you have cognitive impairments [...], yes, you will be more prone to dark patterns because it is harder to process the information that is in front of you, basically because your attention capabilities and decision-making capabilities are reduced” (P3).
- *People who are financially vulnerable.* People with lower income might fall more for dark patterns, especially dark patterns that build on manipulating the price: “People

who are poorer are often more enticed by lower prices than people with more money” (P5).

- *People from specific age groups.* Some experts mentioned age-related vulnerable groups, for example, older generations or children: “Children, obviously, are much more vulnerable to anything like this because they don’t have the critical problem-solving reasoning that is required” (P4).
- *People with low numeracy.* Such individuals might be significantly affected by tricky pricing strategies. For instance, this effect might be prominent in a country with a significant volume of people with low numeracy, as explained in the example provided by the experts. “[T]he ability to interpret relatively simple or not simple but normal pricing information and financial commitments [...] we think that maybe half of the [working age] population have a problem with that. And we think that might particularly overlap with [...] reference pricing” (P7).

Essentially, all our interviewees agreed that the concept of vulnerability in the digital market context should be redefined. This is because, here, all consumers can be vulnerable. One of our experts divided vulnerability into two categories: the first category relates to *personal characteristics*, such as old age and lifestyle, which are closely associated with some of the traditionally vulnerable groups mentioned above, and the second category is *context-dependent vulnerability*, which was widely discussed by our interviewees and is described below.

Context-dependent vulnerability becomes an issue when companies collect extensive amounts of information about an individual, because such information might be used to target vulnerable qualities specific to that individual. “Because the company knows a lot about you, they could know your weaknesses, your disabilities, or that you are currently in a bad place in your life, or whatever, and they could target you with marketing or offers that are directly linked to where you are at the moment, e.g., after divorce or whatever.” (P1).

Multiple factors make consumers vulnerable, as described by P3: “If you are in a hurry, and, I don’t know, say you have done a long shift at work and you are on the bus on the way home, and you have been arguing with your spouse, like everything is against you that day, and also the physical environment is against you, you have essentially put yourself in a vulnerable situation. So you can be contextually vulnerable [...]”. Such context-dependent vulnerability creates an excellent opportunity for unethical business practices such as dark patterns. It enables companies to, instead of targeting one particular group of people, potentially target all consumers. Thus, contextual factors, such as time pressure, and psychological aspects, e.g., mood changes and stress, make people temporarily vulnerable. “I think that for a few days or maybe a week will make me more vulnerable. [...] these kinds of things, whether it is physical events, psychological events, life events, all of those kinds of things day-to-day, week-to-week, make people vulnerable” (P8). Within such a temporary window, customized tricks could be directed at the consumer.

Access to technologies and increasing numbers of online services (and, consequently, transactions) seem to blur the borders between traditionally vulnerable and non-

vulnerable groups. As pointed out by one of our experts, “[...] the digital environment makes that different because we are no longer segregated into environments that are almost entirely defined by class, or by income, or by whatever. Now, because it is known all the time where we are, we can be contacted all the time.”

4.4 Measures to Tackle Barriers

This research’s main objective was to identify ways to overcome the barriers that result from designs that are detrimental to consumers, in terms of both dark patterns and obstacles to understanding and acting on T&Cs. To this end, we analyzed the interview data and used the intervention functions and policy categories defined in the behavior change wheel to identify potential barriers in current policies and interventions and how to reduce them. The main finding was that consumers’ behavior can be changed indirectly (through changes in companies’ behavior or enforcement) and directly (through consumer behavior changes).

4.4.1 Indirect changes to consumers behavior

Policy categories. Overall, most of our experts focused on policy-making and emphasized on issues related to *regulation* and *legislation*. Some of them also mentioned issues related to *guidelines* and *service provision*.

Regulation & Legislation

- *Lack of rules.* One of the barriers related to the current regulations is the lack of rules about auditing and sharing audit logs. According to our experts, design process documentation could be one of the measures to prevent manipulative designs. As explained by P3, there is a need for “[an] audit in which they [companies] go about the decision making for things like how they write their T&Cs, and the steps they have gone through.” Such an approach could improve transparency, and the documentation could be used to prove whether companies have done everything to ensure that their designs do not have manipulative effects on consumers.

The downside of regulations, especially regulations that would require higher transparency, particularly with regard to the design process, is that it might lead to market imbalance. This is because large companies would have an advantage over smaller companies that might not have the resources to comply with these rules and produce detailed documentation or perform the testing needed to estimate whether the design has a detrimental effect on consumers. Hence, another proposition is to ensure that the related rules are specific to company size or industry type.

- *Lack of research-informed regulations.* Another barrier that surfaced in the interviews is that the current regulations are not sufficiently informed by research. For instance, there is not enough information about the harmfulness of different manipulative designs: that is, there is not enough research that quantifies harm. Moreover, even if research on consumer detriment caused by dark patterns or inappropriate representations of T&Cs exists, it is not always searched for or used as a backbone for new regulations. Further, legislation is not always based on research, but it should be based on evidence, mostly empirical evidence: “if you have evidence of harm and market failure, I think things need to be banned” (P4).

- *Lack of outcome-based strategies.* The most prominent approach recommended by experts is an outcome-based strategy to regulate the market. In some countries, a typical way of regulating markets is by mandating that companies disclose a particular type of information, and regulatory bodies, such as consumer agencies, provide more details about such information: “it must be clear, it must be placed in a relevant part of the product page [...] then if the company complies, then everything is fine” (P5). However, such an approach could be improved by mandating that companies provide certain information without explaining how this should be done. Instead, companies could be required to make the information displays effective. This would change the regulator’s role, and they would be responsible for determining whether consumers understand the terms for use of the product or service, instead of asking companies whether they have fulfilled their obligations. “So, it is sort of outcome-based strategy; instead of providing inputs, we specify outputs that must be there, and then we let the firms to fulfill those outputs to the best of their ability” (P5). However, such an approach could increase the current issues related to law surveillance that consumer agencies face. For instance, in the case of the Swedish consumer agency, “We have new measures. We can test purchases and not tell companies that we are from the agency. The problem is that we do not have a Swedish personal number. [...] And this is an obstacle for us because we cannot get a fake personal number and, therefore, we cannot act fully undercover” (P6). Potential solutions to such problems were suggested by another expert: “You are doing the test for an average representative sample of consumers and seeing whether firms are meeting the standard” (P8). Thus, it is possible to use real consumers’ panels and apply different anonymization methods to analyze the data and ensure integrity.
- *Lack of dialogue.* One of the barriers is that there is not enough dialogue between companies and regulators. Our experts believe that a dialogue between companies and consumer authorities could be a good start in identifying the border between nudging and manipulating consumers. “As an enforcer, nearly all the time you feel that you are one step behind. But using the dialogue method at least would allow getting hold of the honest traders” (P2).
- *Lack of prohibited terms.* Legislation should define prohibited terms for T&Cs, and such key information must be communicated to consumers while making sure that only useful information is provided. Some experts even suggested that consent is not always applicable: “I just think that there is a need for a movement away from this idea of consent. Because I think, for certain things it makes sense, but for others, it does not” (P4).

Guidelines

- *Lack of documentation checklists.* There is no concise information on how to document the design process. Therefore, companies should be provided with checklists containing requirements for the documentation of their design process and, consequently, be required to “prove that they are honest” (P3).
- *Lack of clear requirements.* There are no clearly defined lists of disclosure requirements: “we have to specify what information the consumer needs, and then, whenever we do that, we also have to specify how” (P5).

- *Lack of guides.* Currently, smaller companies might not be able to implement extensive research to support their design choices. The larger the company, the more resources it will have to test and identify potentially misleading designs or inappropriate T&Cs. Thus, small companies might be at a disadvantage. Therefore, guides, templates, or design patterns could help companies to comply with the legal requirements. “[...] Producing this kind of best practice guides, and design patterns that companies can use, at least as a starting point. [...] So maybe those two things go hand in hand, is that you need suggested, recommended check-out flows, as well as outcome-based regulation” (P8).

Service provision

- *Lack of government-approved online services.* There are no online services that provide companies with standardized information. One way to improve the display of information, particularly in T&Cs, could be to enable companies to contact the appropriate government body and receive a list of common factors related to the transaction, customized to a given industry type. Such information would be in a preset form: “the government sends them [companies] URL or an HTML webpage that has a government standard check-out page, with all information in it” (P8).
- *Lack of enforcement tools.* Currently, not all of the dark patterns or T&Cs can be analyzed with automatic tools, and this potentially affects the effectiveness of enforcement. “We must get better tools for picking up those data manipulations or similar. We really have to put ourselves in the consumer’s shoes” (P6). Different tools could help enforcers, e.g., automatic tools for analyzing the texts of T&Cs that could be used by lawyers or tools for identifying misleading text-based and visual designs.

Intervention functions. The experts mentioned only a few intervention functions of the behavioral change wheel: *education*, *coercion*, *modeling*, and *training*.

- *Education.* There is no code of ethics implemented in the educational systems that could inform designers or other people who play an active role in technology design. Such education could improve the current situation and potentially reduce both intentional and unintentional “dark” designs. However, our experts did not perceive education as the most robust way of making any changes: “I think codes of ethics don’t work [...]. They are useful in education systems, to teach designers what should matter or to teach people in the industry but ultimately you need to have laws that impact companies commercially, that have a financial impact on them” (P3).
- *Coercion.* There is a lack of significant punishment for the use of misleading design. Some of the experts proposed that companies might change their behavior only when they are confronted with punishment or costs, for example, big fines for not complying with rules. The fines should be considerable enough to bring about change, unlike fines from the past, for instance, the fines placed on LinkedIn for address book leeching that were relatively small considering the actual company profits.
- *Modeling.* The coercion intervention could be used as an exemplar of bad designs and their consequences. “I think some really well-known cases, where the worst of any companies are really nailed to the wall for doing bad things, would require companies to really take the law into action. I think that would help wake up a lot of [...] businesses. [...] I think some really public cases of enforcement, of the worst

offenders, would really help the industry as a whole” (P4). This might be particularly important from the perspective of smaller companies, which may not have a team of lawyers but may base their design decisions on their competitors’ design decisions presented in news outlets.

- *Training*. There seems to be insufficient training in the enforcement agencies, particularly in countries where there is already well-developed legislation about consumer rights and protection. In some cases, enforcement agencies might lack personnel who have a comprehensive view of the digital market. Therefore, appropriate training and education about the digital market’s fast-changing landscape could be beneficial for enforcement agencies.

4.4.2 Direct changes to consumers behavior

In our interviews, direct change in consumers’ behavior was not widely discussed. This implies that consumers should not be overloaded, in contrast to some other areas such as the privacy field, where the user must understand and learn about how their information might be used. In digital markets, consumers’ informed decision-making responsibility seems to have shifted more towards policymakers, regulators, and service providers.

Only on a few occasions, our experts mentioned consumers’ education and awareness. “Awareness-raising is, of course, something very important as well” (P1). However, it is hard to raise awareness, for instance, about second-generation dark patterns, as they are hard to detect. Still, *communication and marketing* as an intervention function might be one of the ways to improve consumers’ decision-making process: “as a consumer agency, we have an obligation to inform consumers, but I am not sure whether they will listen to us. [...] Maybe we could provide some information on our website or on the telephone” (P6).

5 Summary and Discussion

In this report, we attempted to shed light on the many co-existing barriers to well-functioning digital markets. Here, we describe the large volume of tricks that online companies may use to exploit consumers' vulnerabilities and intentionally guide them towards the company's preferred decisional outcome. In this section, first, we summarize and discuss our findings in terms of (1) the different categories and harmfulness of dark patterns, (2) barriers to understanding and acting on T&Cs, (3) consumers' vulnerability, and (4) measures that should be applied to ensure a well-functioning market. We follow this summary with a short discussion about the ethical aspects of malicious design and the effectiveness of dark patterns.

5.1 Dark Patterns

We simplified the categorization of dark patterns with the help of the COM-B model, according to which they were categorized under two main clusters. Moreover, the results from the experts' interviews raise questions about dark patterns, particularly about their harmfulness. We elaborate on these points in the sections below.

5.1.1 Categorization

Our simplified categorization of dark patterns comprises two clusters: those that affect *active choice* and those that affect *information comprehension*. The first cluster contains dark patterns that may have a direct impact on consumers' actions and includes *Bad defaults*, *Toying with emotions*, *Activity message*, *Confirmshaming*, *False hierarchy*, *Hidden cost*, *Bait and switch*, *Price comparison prevention*, *False urgency*, *Rewards and punishment*, *Testimonials*, *Intermediate currency*, *Social pyramid*, *Forced continuity*, and *Sneak into basket*.

Interestingly, most patterns that affected active choice seem to influence consumers' motivation, and the majority have a direct relationship with either the price of the product or service or payment and delivery. Based on the expert responses regarding the severity of harm and consumer detriment resulting from dark patterns, it is clear that the active choice patterns align with patterns that affect consumers' economic loss. They prey on consumers' weaknesses to manipulate their choice towards decisions that are most beneficial for companies but not necessarily the most optimal for consumers. Dark patterns causing economic loss were among the most prominently discussed patterns by our experts, in terms of the severity of consumer detriment these patterns are associated with. These dark patterns are more traditional, first-generation patterns that, perhaps, might be detectable by enforcers or even by consumers and, consequentially, might be easier to regulate.

The second category of dark patterns—those affecting comprehension—are designs that prevent the understanding of presented information and influence consumers' awareness of decisional outcomes. This group contains the dark patterns *Social pyramid*, *False action and timing*, *Address book leeching*, *Friend spam*, *Trick questions*, *Hidden legalese stipulations*, *Bad defaults*, *Toying with emotions*, *Forced continuity*, *Forced enrollment*, *Immortal accounts*, *Hard to cancel/roach motel*, *Privacy Zuckering*, and *Disguised ads*. It seems that most of the dark patterns from

this category have a direct relationship with either T&Cs or consumers' rights after purchase. Additionally, these dark patterns have a more evenly distributed effect on the three drivers of behavior: capability, opportunity, and motivation. However, within this category, capability seems to be most affected (with eight dark patterns displaying this tendency). Additionally, the dark patterns affecting comprehension are more harmful and were mostly classified by our experts as second-generation dark patterns. Thus, these patterns are more insidious and difficult to detect by both consumers and regulators. Consequentially, they are harder to regulate and prohibit.

5.1.2 Harmfulness of dark patterns

As mentioned above, our experts perceived the second-generation patterns to be the most harmful. The harmfulness of these dark patterns seems to lie in the difficulty involved in detecting them, both by consumers and regulators. Second-generation dark patterns are linked to patterns influencing privacy, and privacy-related dark patterns are based on extensive data collection that enables more accurate targeting of an individual's weakness. This close relationship between second-generation dark patterns and data collection practices indicates that data protection and privacy have become more relevant to the digital market than ever before. Given this situation, regulatory bodies must collaborate in order to regulate and maintain a well-functioning market. For instance, consumer agencies should engage in sustained dialogues with companies as well as the appropriate data protection agencies to ensure appropriate data flow.

The second most harmful dark patterns identified by our experts are dark patterns that cause economic loss. These patterns are easier to detect and regulate. However, they are still damaging to the consumer, especially if companies exploit a target consumer's temporary or permanent vulnerability. Other dark patterns perceived as harmful by the expert were those which restrain, i.e., severely restrict, consumers' autonomy, by tangibly affecting their freedom of choice and ability to "shop around." Dark patterns that cause economic loss and restrain might be easier to identify automatically and regulate, for example, by developing intelligent tools based on regulation. Similarly, their damaging effects might be less difficult to understand and quantify, unlike the effects of second-generation patterns that are hidden in the complex nets of data interdependencies.

5.2 Terms & Conditions

By investigating past studies and considering our experts' responses, we attempted to identify the main barriers that prevent consumers from understanding and acting on T&Cs. We identified two main groups of barriers. The first group is related to the content of T&Cs, that is, barriers that make it cumbersome to understand T&Cs and potentially cause a negative attitude towards documents, accelerating the existing habitual effect of automatic acceptance. Under this group, we include the following barriers: *length*, *information overload*, *complexity*, and *unusability*. The second group of barriers are related directly to the visual design of T&Cs, as their visual design frequently lacks usability and usefulness. In general, according to our experts, T&Cs in their current form are inappropriate, and are probably one of the least effective ways of information

disclosure. In this category, we list *legal defensiveness* (e.g., use of the tick box), *lack of clearly defined patterns*, *secondary importance*, and *lack of actionability*.

Under the two groups mentioned above, we have listed some sub-categories of barriers and mapped them to the COM-B framework. From the mapping, it seems that the main categories of barriers to understanding and acting on T&Cs mostly affect capability by preventing consumers from gaining knowledge. Only one of the main categories—*unusable*—seems to influence motivation through its effects on habitual processes (i.e., making decisions as a result of routine and habit), and conscious decision making. An interesting finding that emerged is that most of the sub-categories of barriers influence opportunity. This is perhaps because the sub-categories are more actionable and easier to notice. For instance, *lack of choice* and *lack of actionability* directly diminish consumers' opportunity to act. On the other hand, only three sub-categories impact motivation—*meaningfulness*, *secondary importance*, and *habituation*.

The results of COM-B mapping indicate that targeting barriers that influence capability through appropriate intervention functions and policy changes could alter the effects of these barriers on opportunity and motivation. According to the behavioral change wheel, to change capability, one should implement intervention functions related to education, enablement, or training. These functions, in turn, are linked to some of the policy categories, as shown in Table 5 and discussed in the next section.

Another interesting theme that surfaced from the experts' responses was the usability of T&Cs. Some suggested that T&Cs are, perhaps, not always needed. People tend to have too many contracts in the digital world, and it is almost impossible to learn the details of all of them. On the other hand, considering the similarities between such agreements, consumers behave somewhat rationally through habitual acceptance of the terms without reading them.

The abovementioned usability of T&Cs and the barriers identified in this research call for more research around the actionability of T&Cs. To the best of our knowledge, not many studies have investigated whether changes in the display of T&Cs could affect consumers' behavior, i.e., whether it is possible to overcome the effects of habituation, etc., through more rational behavior. Such research requires field studies or longitudinal experiments. The existing research, which mostly uses traditional experiments, focuses on changes in the design of T&Cs and the effects of different designs on understanding of and engagement with contracts. Instead, future work could examine how people enact such design changes and whether it impacts their purchasing decisions.

5.3 A Well-functioning Digital Market

Our research's main goal was to present an overview of barriers to the well-functioning digital market by focusing on consumer detriment. Accordingly, an important objective was to assess what measures could be applied to reduce consumer detriment. To ensure that the digital market is well-functioning, policy makers and regulators should consider the effects of malicious designs on consumers. Considering that such designs prey on consumers' weaknesses, it is important to understand who the vulnerable digital consumer is.

According to our findings, all consumers can be vulnerable. That is, the traditional assumptions about vulnerable groups of consumers do not necessarily apply to the digital space. Nonetheless, some of the traditionally perceived vulnerable groups will certainly be affected more than others by manipulative designs, whether they are dark patterns or ineffective presentations of T&Cs. However, our experts agreed that people with cognitive impairments, financially vulnerable persons, children, older generations, or people with low numeracy could be potentially more affected. Despite this, policymakers must consider that everyone may be vulnerable because of how the technology works. All the information collected about individuals, including personal information related to their economic, mental, and physical status, is available to exploit and target these individuals. Modern technologies, such as different devices equipped with multiple sensors, and various applications used by a specific individual, enable the collection and aggregation of data from many sources. Using advanced algorithms, companies can detect, in some cases in real time, the different states that a particular individual is in and target that person with “customized” dark patterns. For instance, recommender technologies might suggest products that may seem necessary at a price that is not optimal based on the consumer’s financial status.

5.3.1 Measures to Ensure a Well-functioning Market

Using the behavior change wheel framework, we attempted to identify measures that should be applied to remove the barriers that currently prevent the market from functioning well. Our experts provided a wide range of recommendations that could be used, and these were mapped to the policy categories and intervention functions. These recommendations apply to companies and regulators, and not to consumers. The only advice provided by experts that was related directly to consumers was education and raising awareness. However, these recommendations may be inefficient because some malicious designs are difficult to detect by consumers.

Table 5: Links between policy categories and intervention functions (blue highlights indicate potential links between policy and interventions identified in the current research).

	Education	Persuasion	Incentivization	Coercion	Training	Restriction	Environmental restructuring	Modeling	Enablement
Communication/ Marketing	√	√	√	√				√	
Guidelines	√	√	√	√	√	√	√		√
Fiscal			√	√	√		√		√
Regulation	√	√	√	√	√	√	√		√
Legislation	√	√	√	√	√	√	√		√
Environmental/ Social planning							√		√
Service provision	√	√	√	√	√			√	√

Overall, drawing on our interview data, we identified four policy categories and four intervention functions that currently create barriers to a well-functioning market. Considering the links between policy categories and intervention functions, policymakers and intervention designers might be able to use this information. Table 5 presents such linkage based on the behavior change wheel, by highlighting the issues identified in the current research.

Policy categories

- Regulation & legislation
 - There is a lack of clearly defined rules related to documenting designs and the development process.
 - The current regulations are not informed by research.
 - Regulations should be customized, e.g., according to industry or company size, to ensure a balanced digital market.
 - There is no outcome-based approach; instead, the current approach is a bottom-up one.
 - There is a lack of collaboration between companies and regulators. Open dialogue is necessary to identify borders between dark patterns and nudging.
 - There is no knowledge transfer between the research and legislation. Legislation should be created based on empirical evidence, by incorporating results from multidisciplinary fields, such as psychology, behavioral science, computer science, and law.
 - There is a lack of clearly defined, precise information that could be disclosed to consumers to ensure their comprehension and empower their decisions.
- Guidelines
 - There are no checklists for the documentation of the design process.
 - There are no guidelines with clearly defined requirements.
 - There is a lack of guides about consumer-empowering designs. Such guides could be used by companies that cannot afford to perform user research independently.
- Service provision
 - There is a lack of government-approved online services. Such services could help with the design process, particularly in the case of smaller companies.
 - There is lack of automated tools for helping regulators identify malicious design.

Intervention functions

- *Coercion*. The fines for not complying with rules are not big enough to deter companies. The imposition of considerably high fines could lead companies to change their behavior.
- *Modeling*. There are no examples of large players that broke the law with a non-compliant design and were accordingly punished. Such exemplars could serve as motivation for other companies to produce compliant designs.
- *Training*. Insufficient training of the enforcement workforce is an obstacle. Enforcement agencies could train their employees to fully understand the risks and effects of malicious designs and stay up-to-date with the rapidly changing digital market.

- *Education*. There is a lack of knowledge about relevant topics, for instance, ethical design. Educating designers and companies about the effects of malicious designs, e.g., by creating codes of ethics for designers, could decrease the use of dark patterns and ineffective disclosures.

5.4 Ethical issues

Some of the dark patterns are either unethical or not compliant with the existing regulations. For instance, *immortal accounts* might violate compliance with the General Data Protection Regulation (GDPR) if applied in a service provided to users located within the European Union. This is because the GDPR requires service providers to ensure that consumers can easily access and delete their accounts and all data associated with the service.

5.4.1 Manipulation through Dark Patterns

Considering the ethical issues, dark patterns and the barriers to understanding T&Cs can be perceived as manipulative and, consequently, influencing the autonomy of consumers' decisions. In the context of nudging, the ethical aspects of influencing behaviors have surfaced. Even if nudging is considered as a behavioral influence aimed to improve people's well-being, there is still a possibility that a particular choice architecture might be manipulative. In the context of digital markets, it is becoming more apparent that the effects of such manipulative designs may reduce consumers' autonomy. Legal scholars have already recognized this as problematic: "digital market manipulation is a problem, if at all because it constitutes a form of persuasion that is dangerous to consumers and society" (p.5, [37]).

There seems to be two ways through which people's decisions might be influenced [37]: first, through a change in the available options, that is, the so-called decision space, and second, through a change in how people understand their options or the so-called internal decision-making process. These two strategies are often applied not necessarily to manipulate, but also to persuade someone (through rational persuasion, which is defined as changing someone's mind by providing reasons to further reflection and evaluation). It seems that some dark patterns might be only persuasive while others could be coercive. However, coercion does not follow these two ways of influencing decisions. Instead, it provides people with irresistible incentives. Moreover, coercion usually works on the premise that all other acceptable alternatives are eliminated, and this can influence healthy and balanced market competition. Manipulation, on the other hand, is contradictory to both persuasion and coercion, as manipulation takes away the possibility of control. Therefore, it violates autonomy by depriving individuals of the authorship of their actions [37], and "feeling manipulated" mostly means that a person does not fully understand why they acted in a particular way or whether their actions were beneficial for themselves or others ("manipulator controls his victim by adjusting her psychological levers") (p. 44, [38]). Thus, such dark patterns manipulate by influencing belief, desire, and emotions (i.e., psychological levers).

Dark patterns and, to some extent, the current presentations of T&Cs, might be manipulative because they integrate the three key elements of manipulative practices [29]. That is, these patterns contain influences that are:

1. Hidden;
2. Exploit cognitive, emotional, or other vulnerabilities related to the decision-making processes;
3. Targeted (i.e., they target a consumer being confronted with a specific choice, in a particular context, where the agent is fully aware).

If every choice that is made is always constrained by external factors, such as social or situational context, why is online manipulation an obstacle to autonomy? This is because manipulative choice architectures challenge two conditions of autonomy: competency (an autonomous person has the competency to deliberate, form intentions, and act upon them) and authenticity (an autonomous person can act based on their values and reasons) [29]. If these two conditions are not satisfied, the choice is no longer ours. That is, the choice is not autonomous, as we no longer know what decisions are being made and why we are making those decisions.

5.4.2 Nomenclature

The other ethical issue related to malicious designs concerns the designers and nomenclature of dark patterns. Previous literature [10] has emphasized that the designers' role is crucial and that they should ensure that design strategies are ethically grounded. It is possible that some of the malicious designs are not intended; that is, they result from a lack of skill or knowledge (so-called anti-patterns). However, exploitative techniques are frequently and deliberately applied in choice architectures to increase profitability. Therefore, a designer should ensure that such practices are at a minimum and that their work does not blindly follow the requirements defined by stakeholders seeking to maximize financial profit.

Unintentional dark patterns and the current findings related to dark patterns, particularly the opinions we gathered from experts, raise questions about terminology. Specifically, it seems dubious to perceive dark patterns only as an obstacle when implemented in the system design. As mentioned by some of our interviewees, dark patterns may provide consumers with useful information. For instance, some of the personalized services might be regarded as beneficial. This is one reason why it might be hard to draw a line between patterns that are "dark" and those that are useful or those that are not intentionally "dark."

Perhaps we should consider renaming dark patterns as, merely, malicious designs or inverted nudges. For instance, based on our findings, it seems that second-generation and first-generation dark patterns are comparable with some categorizations of nudges. One could say that first-generation dark patterns are transparent interventions and influence behavior and, therefore, affect the automatic mode of processing. On the other hand, second-generation dark patterns are non-transparent interventions that manipulate choice by influencing the reflective mode of thinking. While not all dark patterns can be accurately categorized under these types, the findings still imply the

presence of a strong bond between the two digital phenomena of nudging and dark patterns.

Theoretically, the definition of dark pattern, which states that a dark pattern intentionally misdirects users into action that is beneficial to the agent, is precise and implies that it should be unproblematic to detect malicious designs. However, to identify such manipulations, it is frequently necessary to assess harm to the consumer. This is difficult, as harm cannot be clearly defined and not all harm is easily measurable, as hinted by some of the experts. It is much easier to measure real damage, such as an economic loss. Yet, it is difficult to identify psychological harm or harm that is not immediate, e.g., harm that resurfaces after a long time due to the extensive data collection. Moreover, not all consumers perceive harm in the same way. Even if financial harm occurs at the time of purchase, in the case of some consumers, buying a product that was not initially desired might eventually be beneficial. It is unlikely that all consumers will be affected in the same way by identical designs because consumers possess unique individualistic characteristics.

The ambiguity about the terminology used in the context of dark patterns needs to be addressed in future research through more large-scale studies and the development of new methods to measure consumers' harm. Such studies could also improve our understanding of what information consumers find useful and how such information disclosures should be regulated to keep consumer detriment at a minimum.

5.5 Effects of Dark Patterns on Companies

There is growing interest in malicious designs; mainly, more researchers are discussing dark patterns and their potential implications on end-users and online service providers. However, there is not much empirical work that has attempted to identify which dark patterns are the most successful or, if we were to rephrase it, which dark patterns can bring the most benefits to the service provider. We found only a couple of studies related to this subject.

In one of the meta-analyses, the authors tried to compare the effects of different nudging practices, both in a digital and physical context [39]. Overall, they found that digital nudging does not differ significantly from nudging in other contexts. However, the analysis showed that approximately one-third of the effect sizes reported in the existing nudging studies were insignificant. Importantly, the default nudges (corresponding to the *bad default* dark patterns) seem to be the most effective.

Another study, which was empirical and longitudinal, investigated the effectiveness of dark patterns and other changes in the UI design [40]. The company that conducted this study collected data from e-commerce websites that implemented A/B testing (comparative testing of different versions of a product to identify which one consumers prefer) of various designs between 2014 and 2017. The results showed that only specific design changes carry the potential to increase revenue per visitor. Unfortunately, the generally classified categories of dark patterns—scarcity, social proof, urgency, abandonment (persuading the user not to leave the site, which indicates abandonment behavior), and product recommendation—seem to be some of the practices that might

influence the user and increase revenue. However, the revenue increase identified in this research was relatively small and ranged from +0.4% to +2.9%.

It seems that on the one hand, dark patterns might cause damage, impact consumers' autonomy, and violate the principles of a fair market, but to some extent, they might be effective. Thus, there is a need to identify which dark patterns cause the most damage and regulate their applicability in online commerce in a way similar to how the GDPR regulates data protection. On the other hand, considering the potentially small benefits that dark patterns could have for companies, educating stakeholders and designers might be a potentially beneficial strategy. However, extensive empirical evidence showing that malicious designs result in low profits would have to be collected. Thus, providing stakeholders with information that such manipulative strategies are not highly profitable and presenting guidelines on attracting consumers to the goods or services in an ethical manner might be an effective way to reduce the use of malicious design in e-commerce services.

6 Bibliography

- [1] Consumers, Health, Agriculture and Food Executive Agency on behalf of Directorate-General for Justice and Consumers, *Operational guidance document on measuring personal consumer detriment*, Report, February, 2017.
- [2] OECD, “Improving Online Disclosures With Behavioural Insights,” *OECD Digital Economy Papers*, no. 269, 2018.
- [3] H. Beales, R. Craswell, and S. C. Salop, “The Efficient Regulation of Consumer Information,” *The Journal of Law and Economics*, vol. 24, no. 3, pp. 491–539, 1981, doi: 10.1086/466997.
- [4] J. Luguri and L. Strahilevitz, “Shining a Light on Dark Patterns,” *SSRN Electronic Journal*, pp. 1–51, 2019, doi: 10.2139/ssrn.3431205.
- [5] K. B. Cornelius, “Zombie contracts, dark patterns of design, and ‘documentisation,’” *Internet Policy Review*, vol. 8, no. 2, pp. 0–25, 2019, doi: 10.14763/2019.2.1412.
- [6] J. S. B. T. Evans and K. E. Stanovich, “Dual-Process Theories of Higher Cognition: Advancing the Debate,” *Perspectives on Psychological Science*, vol. 8, no. 3, pp. 223–241, 2013, doi: 10.1177/1745691612460685.
- [7] D. Kahneman, “A Perspective on Judgment and Choice,” *American Psychologist*, vol. 3, no. 4, pp. 7–18, 2003, doi: 10.1037/0003-066X.58.9.697.
- [8] R. Thaler and C. Sunstein, *Nudge. Improving Decisions About Health, Wealth, and Happiness*. 2008.
- [9] G. Gigerenzer and W. Gaissmaier, “Heuristic Decision Making,” *Annual Review of Psychology*, vol. 62, no. 1, pp. 451–482, 2011, doi: 10.1146/annurev-psych-120709-145346.
- [10] C. M. Gray, Y. Kou, B. Battles, J. Hoggatt, and A. L. Toombs, “The dark (patterns) side of UX design,” *Conference on Human Factors in Computing Systems - Proceedings*, vol. 2018-April, no. April, 2018, doi: 10.1145/3173574.3174108.
- [11] B. J. Fogg, “A Behavior Model for Persuasive Design,” 2003, doi: 10.1111/j.1751-1097.1992.tb02203.x.
- [12] A. Kitkowska, Y. Shulman, L. A. Martucci, and E. Wästlund, “Psychological Effects and Their Role in Online Privacy Interactions : A Review,” *IEEE Access*, vol. 8, 2020.
- [13] A. Caraban, E. Karapanos, D. Gonçalves, and P. Campos, “23 Ways to Nudge: A review of technology-mediated nudging in human-computer interaction,” *Conference on Human Factors in Computing Systems - Proceedings*, pp. 1–15, 2019, doi: 10.1145/3290605.3300733.

- [14] F. Slovic, "The Affect Heuristic," in *Heuristics and Biases; The Psychology of Intuitive Judgement*, Cambridge University Press, 2002, pp. 397–420.
- [15] A. Tversky and D. Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science*, vol. 185, no. 4157, pp. 1124–1131, 1974, doi: 10.1017/cbo9780511809477.002.
- [16] D. Schneider, S. Lins, A. Benlian, and A. Sunyaev, "Nudging Users Into Online Verification : The Case of Carsharing Platforms," *Thirty Eighth International Conference on Information Systems, South Korea 2017*, pp. 1–20, 2017.
- [17] T. Mirsch, C. Lehrer, and R. Jung, "Digital Nudging: Altering User Behavior in Digital Environments," *Proceedings of 13th International Conference on Wirtschaftsinformatik*, pp. 634–648, 2017.
- [18] J. P. Simmons, L. D. Nelson, and U. Simonsohn, "False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant," *Psychological Science*, vol. 22, no. 11, pp. 1359–1366, 2011, doi: 10.1177/0956797611417632.
- [19] D. Kahneman, J. L. Knetsch, and R. H. Thaler, "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias," *Journal of Economic Perspectives*, vol. 5, no. 1, pp. 193–206, 1991, doi: 10.1257/jep.5.1.193.
- [20] R. H. Thaler and S. Benartzi, "Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy*, vol. 112, no. S1, pp. S164–S187, 2004, doi: 10.1086/380085.
- [21] R. F. Baumeister and B. J. Bushman, *Social Psychology and Human Nature*. 2013.
- [22] D. Ariely, A. Bracha, and S. Meier, "Doing Good or Doing Well ? Image Motivation and Monetary Incentives in Behaving Prosocially," vol. 99, no. 1, pp. 544–555, 2007.
- [23] D. Halpern, I. Vlaev, P. Dolan, M. Hallsworth, D. King, and R. Metcalfe, "Influencing behaviour: The mindspace way," *Journal of Economic Psychology*, vol. 33, no. 1, pp. 264–277, 2011, doi: 10.1016/j.joep.2011.10.009.
- [24] D. Kahneman and S. Frederick, "Representativeness revisited: Attribute substitution in intuitive judgment," *Heuristics of Intuitive Judgment: Extensions and Applications*, no. January 2002, pp. 1–30, 2002, doi: 10.1038/2251090a0.
- [25] C. J. Anderson, "The psychology of doing nothing: Forms of decision avoidance result from reason and emotion," *Psychological Bulletin*, vol. 129, no. 1, pp. 139–167, 2003, doi: 10.1037/0033-2909.129.1.139.

- [26] The Behavioural Insights Team, “Improving consumer comprehension of contractual terms and privacy policies online: Literature review,” no. August, pp. 1–30, 2018, [Online].
- [27] S. Michie, M. M. van Stralen, and R. West, “The behavior change wheel: A new method for characterising and designing behavior change interventions,” *Implementation Science*, vol. 6, no. 42, 2011, doi: 10.1001/archderm.1985.01660070119033.
- [28] “Dark Patterns - About us.” <https://darkpatterns.org/about-us.html> (accessed Jan. 12, 2021).
- [29] C. Bösch, B. Erb, F. Kargl, H. Kopp, and S. Pfattheicher, “Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns,” *Proceedings on Privacy Enhancing Technologies*, vol. 2016, no. 4, pp. 237–254, 2016, doi: 10.1515/popets-2016-0038.
- [30] A. Mathur, G. Acar, M. Friedman, E. Lucherini, J. Mayer, M. Chetty, A. Narayanan, “Dark Patterns at Scale : Findings from a Crawl of 11K Shopping Websites,” *Proceedings ACM Human-Computer Interaction 3, CSCW*, Article 81, 2019.
- [31] N. Steinfeld, “‘I agree to the terms and conditions’: (How) do users read privacy policies online? An eye-tracking experiment,” *Computers in Human Behavior*, vol. 55, pp. 992–1000, 2016, doi: 10.1016/j.chb.2015.09.038.
- [32] M. Lippi, P. Palka, G. Contissa, F. Lagioia, H.-W. Micklitz, G. Sartor, P. Torroni, “CLAUDETTE: an automated detector of potentially unfair clauses in online terms of service,” *Artificial Intelligence and Law*, vol. 27, no. 2, pp. 117–139, 2019, doi: 10.1007/s10506-019-09243-2.
- [33] E. Luger, S. Moran, and T. Rodden, “Consent for All: Revealing the Hidden Complexity of Terms and Conditions,” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, p. 2687, 2013, doi: 10.1145/2470654.2481371.
- [34] T. F. Waddell, J. R. Auriemma, and S. S. Sundar, “Make it Simple, or Force Users to Read? Paraphrased Design Improves Comprehension of End User License Agreements,” in *Proceedings of Conference on Human Factors in Computing Systems - CHI 2016*, 2016, pp. 0–4.
- [35] M. Elshout, M. Elsen, J. Leenheer, M. Loos, and J. Luzak, *Study on consumers’ attitudes towards Terms and Conditions (T&Cs): final report*. 2016.
- [36] Danish Competition and Consumer Authority, “Improving the effectiveness of terms and conditions in online trade,” pp. 1–6, 2018.

- [37] D. Susser, B. Roessler, and H. Nissenbaum, "Online Manipulation: Hidden Influences in a Digital World," vol. 53, no. 9, pp. 1689–1699, 2013, doi: 10.1017/CBO9781107415324.004.
- [38] R. Noggle, "Manipulative actions : A conceptual and moral analysis," *American philosophical quarterly (Oxford)*, vol. 33, no. 1, pp. 43–55, 1996, doi: 10.2307/20009846.
- [39] D. Hummel and A. Maedche, "How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies," *Journal of Behavioral and Experimental Economics*, vol. 80, no. September 2018, pp. 47–58, 2019, doi: 10.1016/j.socec.2019.03.005.
- [40] W. Browne and M. S. Jones, "What works in e-commerce - a meta-analysis of 6700 online experiments," *Qubit Digital Ltd*, pp. 1–21, 2017.
- [41] Behavioral Insights Team, "Best Practice Guide. Improving consumer understanding of contractual terms and privacy policies: evidence-based actions for businesses.," 2019.
- [42] C. Cara, "Dark Patterns in the Media: a Systematic Review," vol. VII, no. 14, pp. 105–113, 2019.
- [43] M. Chromik, S. T. Völkel, M. Eiband, and D. Buschek, "Dark patterns of explainability, transparency, and user control for intelligent systems," *CEUR Workshop Proceedings*, vol. 2327, 2019.
- [44] G. Conti and E. Sobiesk, "Malicious interface design: Exploiting the user," *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, no. April 2010, pp. 271–280, 2010, doi: 10.1145/1772690.1772719.
- [45] ForbrukerRadet, "Deceived By Design," 2018. [Online]. Available: <https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf>.
- [46] A. Narayanan, A. Mathur, M. Chetty, and M. Kshirsagar, "Dark patterns: Past, Present, and Future. The Evolution of Tricky User Interfaces," *Communications of the ACM*, vol. 63, no. 9, pp. 42–47, 2020, doi: 10.1145/3397884.
- [47] M. Nouwens, I. Liccardi, M. Veale, D. Karger, and L. Kagal, "Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence," 2020, doi: 10.1145/3313831.3376321.
- [48] A. Stibe, K. Pangbourne, S. Wells, B. Cugelman, and A.-K. K. Christensen, "Uncovering Dark Patterns in Persuasive Technology," *Persuasive Technology*, pp. 1–4, 2018.

- [49] A. E. Waldman, “Cognitive biases, dark patterns, and the ‘privacy paradox,’” *Current Opinion in Psychology*, vol. 31, pp. 105–109, 2020, doi: 10.1016/j.copsy.2019.08.025.
- [50] F. Westin and S. Chiasson, “Opt out of privacy or ‘go home’: Understanding reluctant privacy behaviours through the fomo-centric design paradigm,” *ACM International Conference Proceeding Series*, pp. 57–67, 2019, doi: 10.1145/3368860.3368865.
- [51] B. Xiao and I. Benbasat, “Product-related deception in E-commerce: A theoretical perspective,” *MIS Quarterly: Management Information Systems*, vol. 35, no. 1, pp. 169–195, 2011, doi: 10.2307/23043494.
- [52] Ş. Özdemir, “Digital nudges and dark patterns: The angels and the archfiends of digital communication,” *Digital Scholarship in the Humanities*, vol. 35, no. 2, pp. 417–428, 2020, doi: 10.1093/lc/fqz014.

Appendix

Appendix A: List of publications identified through the systematic literature review	2
Appendix B: Additional information extracted about the identified dark patterns	4
Appendix C: Definitions of previous categories applicable to the identified dark patterns ...	9

Appendix A: List of publications identified through the systematic literature review

1. Behavioral Insights Team. (2019). *Best Practice Guide. Improving consumer understanding of contractual terms and privacy policies: evidence-based actions for businesses* [41].
2. Bösch, C., Erb, B., Kargl, F., Kopp, H., & Pfattheicher, S. (2016). Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns. *Proceedings on Privacy Enhancing Technologies, 2016* [29].
3. Cara, C. (2019). *Dark Patterns in the Media: a Systematic Review. VII(14)*, 105–113 [42].
4. Chromik, M., Völkel, S. T., Eiband, M., & Buschek, D. (2019). Dark patterns of explainability, transparency, and user control for intelligent systems. *CEUR Workshop Proceedings* [43].
5. Conti, G., & Sobiesk, E. (2010). Malicious interface design: Exploiting the user. *Proceedings of the 19th International Conference on World Wide Web, WWW '10* [44].
6. Cornelius, K. B. (2019). Zombie contracts, dark patterns of design, and 'documentisation.' *Internet Policy Review*, 8(2), 0–25 [5].
7. ForbrukerRadet. (2018). *Deceived By Design* [45].
8. Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The dark (patterns) side of UX design. *Conference on Human Factors in Computing Systems* [10].
9. Luguri, J., & Strahilevitz, L. (2019). Shining a Light on Dark Patterns. *SSRN Electronic Journal* [4].
10. Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale: Findings from a crawl of 11K shopping websites. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW) [46].
11. Narayanan, A., Mathur, A., Chetty, M., & Kshirsagar, M. (2020). Dark patterns: Past, Present, and Future. The Evolution of Tricky User Interfaces. *Communications of the ACM*, 63(9), 42–47 [46].
12. Nouwens, M., Liccardi, I., Veale, M., Karger, D., & Kagal, L. (2020). *Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence* [47].
13. OECD. (2018). Improving Online Disclosures with Behavioural Insights. *OECD Digital Economy Papers*, (269) [2].
14. Stibe, A., Pangbourne, K., Wells, S., Cugelman, B., & Christensen, A.-K. K. (2018). Uncovering Dark Patterns in Persuasive Technology. *Persuasive Technology* [48].
15. The Behavioural Insights Team. (2018). *Improving consumer comprehension of contractual terms and privacy policies online: Literature review. (August)*, 1–30 [26].
16. Waldman, A. E. (2020). Cognitive biases, dark patterns, and the 'privacy paradox.' *Current Opinion in Psychology*, 31, 105–109 [49].
17. Westin, F., & Chiasson, S. (2019). Opt out of privacy or "go home": Understanding reluctant privacy behaviours through the fomo-centric design paradigm. *ACM International Conference Proceeding Series* [50].

18. Xiao, B., & Benbasat, I. (2011). Product-related deception in E-commerce: A theoretical perspective. *MIS Quarterly: Management Information Systems* [51].
19. Özdemir, Ş. (2020). Digital nudges and dark patterns: The angels and the archfiends of digital communication. *Digital Scholarship in the Humanities* [52].

Appendix B: Additional information extracted about the identified dark patterns

Name		
Related patterns	Context	Psychological effects (See Table 1, p. 14)
Previous categorization (See Appendix C, p. 9)		
Activity message		
Testimonials, false urgency	These patterns could be used in mobile and web e-commerce applications, particularly in service-oriented applications related to tourism, but they are not limited to these applications.	The bandwagon effect, social norms, framing, and affect heuristic
Social proof		
Address book leeching		
Friend spam	This pattern is more likely to be used in mobile applications that rely on access to contacts stored on the device.	Because this dark pattern is not visible to the consumer, it does not exploit any psychological vulnerabilities.
Deny, obscure, forced action		
Bad defaults		
Bait and switch, forced registration, forced continuity, forced enrolment, hidden cost	This pattern could be used in both mobile and web e-commerce environments and applications that provide consumers with interactive choices.	Asymmetric information, status quo, default effect
Covert, asymmetric, misdirection, obscure		
Bait and switch		
Disguised ads	This pattern could be used in any mobile or web e-commerce application.	Mere exposure, affect heuristic
Sneaking, misdirection, interface interference		
Confirmshaming		
Hidden information, bad defaults, toying with emotion, trick questions, disguised ads	This pattern could be used in mobile and web e-commerce applications that contain any actionable activities, i.e., when consumers are confronted with a direct choice.	Affect heuristic, framing, and asymmetric information
Interface interference		

Disguised ads		
Hidden information, pre-selection, confirmshaming	This pattern could be used in both web and mobile e-commerce applications. It is disguised during web navigation as buttons or similar features.	Because consumers are not aware of such ads, this dark pattern does not directly exploit any psychological effects. However, it relies on an automatic mode of processing which assumes that consumers will rely on quick information processing and click on the ad.
Interface interference		
False hierarchy		
Bad defaults	This pattern could be used in both mobile and web e-commerce applications in the form of user interface design elements that are more prominent than others, e.g., graying out of certain clickable buttons to make other options more prominent. It could also be applied through non-visual design elements, such as content with direct recommendations.	Anchoring, framing, (possibly also) image motivation
Interface interference		
False urgency		
Forced action and timing	This pattern could be used in e-commerce websites or mobile applications during interactions related to product or service selection.	Framing, availability heuristic (perhaps), scarcity bias
Urgency, just annoying, about sales		
Forced action and timing		
False urgency	This pattern is mostly used in mobile applications; here, consumers might be on the move while interacting with the application.	Default effect, framing, information asymmetry, status quo. Overall, this pattern relies on an automatic mode of decision making in which the decision is constrained by external factors, such as situational context or time pressure.
Forced action		
Forced continuity		
Hidden cost, sneak into basket, bait and switch	This is applicable on the web, usually in the context of services that offer subscriptions.	Asymmetric information, framing
Sneaking, need official regulation, about sales		

Forced enrollment		
Forced continuity, immortal accounts, and bad defaults	This pattern could be applied in any online service that offers account creation.	Instant gratification, asymmetric information
Forced action, Maximize		
Friend spam		
Privacy Zuckering	This could be used in any application that uses social logins or email permissions.	No psychological effects are exploited because consumers are not aware of this pattern.
Forced action, need official regulation, about time spent on a product or service		
Hard to cancel/Roach motel		
Immortal accounts	This could be used in any context that requires the user to create an account, including mobile and web-based e-commerce applications, or in instances of opt-out options.	Information asymmetry
Obstruction, about sales, moderately bad		
Hidden cost		
Sneak into the basket, Hidden subscription	This could be used in both mobile and web applications for the purchase of goods or services.	Affect heuristic, default effect, status quo. Overall, this pattern relies on the automatic mode of processing, under the assumption that consumers will not contemplate on decisions for too long and, instead, decide quickly and impulsively. This reduces the possibility that they will notice the dark pattern.
Sneaking, about sales, moderately bad		
Hidden legalese stipulations		
Privacy Zuckering	Online e-commerce services always contain terms and conditions that explain consumers' rights and protect the company and consumers legally.	The likelihood of consumers missing the opportunity to comprehend all the details in their entirety increases the probability of automated processing.
Obscure		
Immortal accounts		
Roach motel	This pattern could be used by any e-commerce platform that offers account creation services.	Deletion of accounts is made difficult through processes that require consumers to put in cognitive effort in the form of reflective information processing.
Obstruction, deny, obscure		

Intermediate currency/Pseudo currency		
N/A	This could be used in mobile or web applications that contain in-app purchases.	Decoupling, information asymmetry
Obstruction, moderately bad, about sales		
Price comparison prevention		
Forced action and timing	This pattern could be used during purchasing activities, both in web and mobile commerce applications.	Information asymmetry. This pattern relies on an automatic mode of processing, under the assumption that consumers will perceive the action as being too cumbersome or lengthy to act upon and, instead, make the purchase based on limited information.
Obstruction, needs official regulation, about sales		
Privacy Zuckering		
Bad defaults	This could be applied in all e-commerce services that contain privacy policies which consumers need to acknowledge.	Default effect, status quo and, in general, automatic mode of thinking (which implies that the consumer is not keen to spend time and cognitive resources on learning about the situation).
Obscure, forced action		
Rewards and punishment		
N/A	This pattern could be used in both mobile and web-based e-commerce applications.	Affect heuristic, image motivation, (perhaps also) messenger effect
N/A		
Sneak into basket		
Bad defaults	This may be used in any e-commerce application.	Default effect, status quo
Sneaking, about sales, needs official regulation		
Social pyramid		
Friend spam	Any e-commerce service, both mobile and web-based, could use this pattern.	Instant gratification
Forced action		
Testimonials		
N/A	This pattern could be used in any application that contains reviews of the product or service.	Bandwagon effect, social norms
Covert, social proof		

Time countdown		
Toying with emotion, low stock	This could be used in any web or mobile-based application that sells products or services.	Scarcity bias, anchoring, loss aversion
Urgency		
Toying with emotions		
Confirmshaming	This pattern could be used in any web or mobile-based e-commerce application. This pattern is more effective when applied to a specific action that the consumer might take.	Affect heuristic
Interface interference, miscellaneous, moderately bad		
Trick questions		
Toying with emotion, bad defaults, hidden information	This could be used in any e-commerce application, particularly those that involve account creation, or during the check-out process.	Default, framing, anchoring
Misdirection, moderately bad, miscellaneous		

Appendix C: Definitions of previous categories applicable to the identified dark patterns

Social proof. The social proof principle implies that people determine whether a behavior is correct or not, based on others' behaviors. Therefore, in a user interface design, this aspect of human nature can be easily exploited to nudge users towards specific actions. Hence, social proof phenomena can be considered as a dark pattern [4].

Sneaking. This category contains dark patterns that aim to misrepresent consumers' actions or hide/delay information, which, if visible, would make the consumer object to such activities [30][10].

Urgency. Dark patterns that belong to this category impose a deadline on purchasing. Thus, they lead consumers to believe that if no action is taken, the consumers will experience a loss ("fear of missing out") [30].

Misdirection. These dark patterns are based on a purposeful design that focuses consumers' attention on one thing to distract their attention from another thing [30].

Scarcity. This category contains dark patterns that highlight low in-stock items, indicate limited availability of a good or service, or show that they are in high demand [30][40].

Obstruction. . This category contains dark patterns that impede "a task flow, making an interaction more difficult than it inherently needs to be with the intent to dissuade an action" [10].

Forced action. Dark patterns from this category lead users to perform a particular action to obtain a specific functionality of the online application or service [10].

Nagging. This category of dark patterns redirect expected functionality persistently, beyond one or more interactions [10].

Interface interference. Through user interface design elements, dark patterns in this category manipulate consumers by making specific actions more prominent than others [10].

Maximize. This category contains dark patterns that aim to collect an inappropriate amount of data [29].

Publish. Dark patterns in this category enable personal information to be available to the public instead of being hidden from plain view [29].

Obscure. Dark patterns in this category prevent consumers from learning how their personal information is collected, stored, and processed [29].

Deny. This category of dark patterns prevents consumers from having control over their information [29].

Konsument
verket · KO

Konsumentverket, Box 48, 651 02 Karlstad
Telefon: 0771-42 33 00. konsumentverket.se