

Barriers to a Well-Functioning Digital Market: Exploring Dark Patterns and How to Overcome Them

Agnieszka Kitkowska
Karlstad University
agnieszka.kitkowska@kau.se

Johan Högberg
Konsumentverket
Johan.Hogberg@konsumentverket.se

Erik Wästlund
Karlstad University
erik.wastlund@kau.se

Abstract

In a well-functioning digital economy, consumers should be able to make autonomous and informed choices, and companies compete fairly. One of the barriers preventing such well-functioning is dark patterns—designs that mislead users into making specific purchase-related choices. In this research, through a qualitative inquiry (expert interviews), we classify dark patterns based on the harmful ways such designs affect the digital market. Moreover, we analyze data using the behavior change framework and illustrate ways to prevent dark patterns and grant consumers greater protection and autonomy. Our exploratory results outline potential solutions policymakers might apply to improve digital market well-functioning.

1. Introduction

A well-functioning digital market and consumer well-being are factors relevant to a sustainable economy, considering sustainability defined as the improvement of human well-being over time, especially in the economic, social, and environmental sense [1].

In the digital market, the end-user (consumer) empowerment could be one way towards “digital sustainability” [2]. The empowered consumers may drive innovation, productivity, and market competition [3, 4]. They strive to reduce uncertainty [3] and to make optimal decisions, understanding their “own preferences and choices available to them” [4]. Consumer empowerment relies partially on consumers themselves (e.g., skills, knowledge) and partially on the protection, regulation, and institutions developed to help consumers. Therefore, consumers must possess sufficient information, opportunities, and autonomy over their decisions while being granted adequate protection and understand how to utilize it. Nevertheless, today’s digital companies frequently set barriers preventing consumer empowerment and, consequently, the market well-functioning.

We study some of such barriers affecting consumer empowerment through detriment. We refer mainly to the *personal detriment*—“the difference between the value that consumers reasonably expected to get from a good or service and the value that they actually get from it, relating to problems experienced by consumers post-purchase” [5]. To some extent, we also consider issues related to *structural detriment*, of which one characteristic is a regulatory failure.

We focus on one type of barrier—dark patterns—user interfaces (UIs) purposefully designed to mislead users. As a result, users cannot follow their preferences and desires and may become subject to manipulations [6]. Over the last decade, dark patterns received much attention in multidisciplinary research, identifying many categories of dark patterns (e.g., [7, 8, 9, 10]). Two objectives of this study extend the knowledge of the phenomena and add to the existing categorization. First, we attempt to investigate the harmfulness of dark patterns. Second, we aim to provide policymakers insights about changes needed to eliminate detrimental effects that they might have on consumers.

Our results imply the need for new categorization of dark patterns, according to how harmful the techniques that different dark patterns rely on are—what we label as *first-generation* and *second-generation* dark patterns. Further, our findings determine what policymakers could do, mainly to provide consumers with autonomous and informed decisions and ensure fair market competition. We systematize the results according to a well-established behavioral change framework, making our findings easier to apply. Additionally, our results highlight the issues related to consumer vulnerability, notably how the meaning of vulnerability in the digital market diverges from the traditional perception of this construct.

2. Background

In the traditional economic sense, the empowered consumer must possess sufficient information enabling optimal choice—as the ultimate goal is to maximize

| Your Basket | | |
|---|--------------------------------|------------|
| Item | Qty | Price |
|  Plant Lorem ipsum dolor sit amet, consectetur Lorem ipsum dolor sit amet, consectetur | <input type="text" value="1"/> | 1200,00 kr |
|  Greetings Card Lorem ipsum | <input type="text" value="1"/> | 150,00 kr |

Figure 1. Sneak into basket dark pattern exploiting status quo effect—consumer will unlikely make any changes to the predefined selection.

benefit [11]. Consequently, sometimes companies are required by law to disclose specific information, e.g., attributes recommended by the OECD, such as the price for the product/service, Terms & Conditions (T&Cs), payment and delivery, and information about consumer rights after purchase [12]. These information disclosures are usually conveyed to the consumer through UI design.

2.1. Information and UI design

Some suggest that there are two ways through which people’s decisions might be influenced [13]. One way is to change the available options presented to the decision-maker in the decision space. Another way is to change how people understand their options by altering their decision-making processes. The two strategies are common in UI designs, not always to manipulate but also to persuade the user (via rational persuasion, the designer aims to change the user’s mind by presenting the arguments for further reflection and evaluation of a decision). However, some designs manipulate the decision process; they usually target users’ psychological vulnerabilities, preying on cognitive shortcomings.

One of such manipulative UI designs is dark patterns—“tricks used in websites and apps that make you do things that you didn’t mean to, like buying or signing up for something” [14]. Here, consumers cannot follow their preferences and desires and become subject to exploitation [6]. Dark patterns are heavily rooted in psychology and behavioral sciences findings, applying knowledge about humans’ reliance on heuristics and biases into the digital space.

2.1.1. Decision-making. Commencing the online purchasing process, consumers usually have preferences about the goods or services that best suit their current needs. Yet, at times, consumers do not align their choices with their initial preferences. Such a gap between attitudes and behaviors is not unique to digital environments. In psychology and behavioral sciences,

such phenomena have been recognized as typical of human behavior. One way to explain it is through the dual-process theories, implying that people make decisions using the two different information processing modes working in parallel. The first mode is Type 1, based on automatic, fast, and uncomplicated information processing. It is responsible for decisions that do not require working memory [16, 17]. The second mode—Type 2—involves complex, slower information processing, access to working memory, and is often considered more rational [16]. Despite the potential benefits of Type 2 processing, people’s daily decisions mostly rely on Type 1 because it requires less cognitive effort. It utilizes mental shortcuts—heuristics enabling quick and efficient problem solving and decisions. One downside of heuristics is that they may result in cognitive biases (i.e., subjective patterns and inclinations towards choices not always the most optimal). In effect, the decisional outcomes might be less beneficial.

Thaler & Sunstein described how such cognitive shortcomings could be applied in choice architecture to improve human well-being, introducing the concept of nudging [18]. A nudge is “any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives” [18]. Some governments successfully applied nudges to affect decision-makers in the physical space; e.g., default heuristic increased organ donations, decreased waiting times, and lowered the number of deaths [19].

2.1.2. Digital design and dark patterns. The concept of nudging is not a stranger to the online environment, and small changes in the UI design may impact consumer preferences and decisions [20, 21]. Moreover, design influence might be more effective if UIs exploit consumers’ psychological vulnerabilities [18]. For example, when UI contains pre-selected boxes triggering the status quo bias (Figure 1) or when the choice presentation triggers primacy or recency effects.

Online companies might abuse nudging [22] to maximize profits while having adverse effects on users (e.g., economic loss, intrusiveness, reduction of autonomy). One example of such abuse is dark patterns. To tackle the dark patterns and clarify their effects, researchers attempted to categorize them [7, 9, 10, 6]. For instance, [6] summarized the existing categorizations as the main eight categories: nagging, social proof, obstruction, sneaking, interface interference, forced action, scarcity, and urgency. In the context of online shopping, based on a data set from about 11K websites and 53K products, [10] established a

| Interventions | Policy categories |
|--|---|
| <i>Education</i> Increasing knowledge and comprehension | <i>Communication</i> Using print, electronic, telephonic, or broadcast media |
| <i>Persuasion</i> Communication inducing positive or negative feelings to stimulate action | <i>Guidelines</i> Creating document that recommend or mandate practice (including changes in the service provision) |
| <i>Incentivisation</i> Creating an expectation of reward | <i>Fiscal</i> Using the tax system to reduce or increase the financial cost |
| <i>Coercion</i> Creating an expectation of punishment or cost | <i>Regulation</i> Establishing rules or principles of behavior or practice |
| <i>Training</i> Imparting skills | <i>Legislation</i> Making or changing laws |
| <i>Restriction</i> Applying rules that reduce the opportunity to engage in the target behavior | <i>Environmental/Social planning</i> Designing and/or controlling the physical or social environment |
| <i>Environmental restructuring</i> Changing the physical or social context | <i>Service provision</i> Delivering a service |
| <i>Modeling</i> Providing example to aspire to or imitate | |
| <i>Enablement</i> Increasing means/reducing barriers to increase capability or opportunity | |

Table 1. Definitions of intervention functions and policy categories (as presented in Michie et al. [15]).

five-dimensional taxonomy of dark patterns: asymmetric, covert, deceptive, hides information, restrictive.

Despite the extensive literature on dark patterns, it remains challenging to distinguish dark patterns from useful designs. Such distinction is essential because some designs genuinely aim to help users, presenting helpful information or supporting informed choices. In an attempt to tackle this issue, building on previous research, [23] proposed to classify dark patterns according to their effects on user's choice: modifying the decision space (asymmetric, restrictive, disparate treatment, and covert dark patterns) and manipulating the information flow (deceptive, and information hiding dark patterns). Further, they proposed that the normative perspective should be considered when assessing the "darkness" of patterns, considering individual and collective welfare, regulatory objective (instrumental perspective), and an individual's autonomy.

2.2. Behavior change

Taking into account online companies and their influence on consumer behavior, in this research, we applied the COM-B (Capability, Opportunity, Motivation-Behavior) framework [15]. This framework has been developed to help in identifying intervention functions and policy categories that might prevent policymakers from disregarding aspects essential to well-designed policies [15]. Although this framework was initially designed and used in the health context [15], it can be applied in other areas [24], e.g., sustainable circular economy [25].

The COM-B model assumes three components that interact and generate behavior: capability, opportunity, and motivation [15]. Capability refers to an individual's psychological and physical capacity to engage in the given activity. Motivation refers to all the brain processes

that stimulate and direct behavior. The opportunity contains all the factors that lie outside of the individual that make the behavior possible. The COM-B recognizes that different intervention types may change one or more of these three components and consequentially influence behavior. The framework's authors further proposed the behavior change wheel (BCW), offering possible interventions that might steer changes [15]. The BCW includes nine intervention functions and seven policy categories, presented in Table 1.

This work aims to understand the harmfulness of dark patterns better. The other goal is to inform policymakers about the potential solutions that could help ensure a well-functioning digital market, reducing consumer detriment. Hence, we raise two research questions:

RQ1: How can we classify the harmfulness of dark patterns?

RQ2: How can we reduce the detrimental effects that dark patterns have on the consumer?

3. Method

Due to the exploratory nature of our research questions and because we aimed to identify ways to overcome the effects of dark patterns, we used expert interviews¹. The expert interview is a qualitative method, based on a topical guide, focusing on the knowledge of the expert from a certain field [27]. In our work, expert interviews enabled obtaining in-depth knowledge from people familiar with the concept of dark patterns, who have a good understanding of the digital market. We gathered different types of knowledge (technical, process,

¹The study was a part of a larger project; here, we report only on the parts of the project related to dark patterns. The interested reader can access our published report describing methodology and findings in [26].

and interpretive), which could be used to explore and ground theoretical concepts [28].

Before the exploratory interviews, we conducted a literature review [26], identifying previous reviews of research on dark patterns (19 articles). Next, we extracted 27 dark patterns that might affect consumers during online purchasing and gathered information related to T&Cs. Based on our literature review, we composed a semi-structured interview study with four open-ended questions. In this paper, we only report answers to three of these questions². First, we asked which dark patterns might be most harmful to the consumer (a) and what measures should be applied to overcome the obstacles caused by dark patterns and disclosures in T&Cs (b). Next, we asked whether specific user groups might be affected more by dark patterns and contemporary presentations of T&Cs (c).

We interviewed eight experts from law, economy, philosophy, and behavioral and cognitive sciences. The interviews were conducted in English (four participants were native speakers, and the rest were fluent in English). Five experts were selected among the participants of a dark patterns webinar. They either participated in the event actively as presenters or passively as the audience (here, we included some pioneers of the dark patterns and dark patterns' researchers). The rest of our experts were employees from the Swedish Consumer Agency—involved in organizing the webinar, having the expertise related to digital markets and consumer rights, and being familiar with the concept of dark patterns. Due to the COVID-19 pandemic, all interviews were conducted online through a video conferencing platform in November 2020. Before the interviews, each participant received informed consent to sign and return to the researcher in charge, agreeing to record the interviews.

During the interview, participants were informed how the interview would look and explained the topic of interest, including the definition of dark patterns. The interviews' duration ranged from 30 minutes to one hour. The interviews were audio-recorded and manually transcribed. We analyzed transcripts through thematic analysis (RQ1) and deductive analysis based on the COM-B framework—Table 1 (RQ2).

4. Results

We first present findings related to the harmfulness of dark patterns, followed by possible ways to overcome

²We do not report answers to question: *T&Cs are often means used to provide consumers with [informed choice enabling] information. What do you think are the main barriers that consumers encounter regarding understanding and taking into account T&Cs?* This is due to space limitations, clarity, and experimental research motivated by this question's answers (not yet published).

the adverse effects that dark patterns might have on the digital economy.

4.1. Harmfulness of dark patterns

To answer RQ1, we analyzed responses to interview questions (a) and (c). Although we asked the experts about dark patterns' harmfulness, their answers deviated towards discussing how different dark patterns work and how and why certain categories are more or less harmful. Here, the analysis revealed two categories of dark patterns—*first-generation* and *second-generation*. Additionally, interviewees have discussed the concept of vulnerability extensively, particularly considering the latter category of dark patterns.

4.1.1. First-generation dark patterns. The *first-generation* dark patterns are designs that are relatively easy to identify both by consumers and by regulators. They are more likely to self-regulate, as described by P4 “[They make] you accept something you are not willing to accept. I think that it is a problem, but it [...] has a tendency to balance itself because people will neglect or decline from revisiting shops that use these too aggressively.”

One of the themes related to the discussion on the harmfulness was that the *first-generation* dark patterns result in economic loss. They rely on behavioral biases and usually induce hidden costs. As P4 described it: “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.” Some commonly known dark patterns were listed in the context of economic loss, e.g., scarcity-based patterns (Figure 2) or hidden cost.

The *first-generation* dark patterns might be easier to identify and consequently easier to regulate. However, that does not mean these dark patterns are benign. Even though some regulations might forbid such designs, the regulatory bodies may struggle to identify economic loss dark patterns because they might be presented to consumers after creating an account, thus limiting regulators' surveillance techniques.

Other harmful *first-generation* dark patterns restrain consumers, e.g., lock consumers in one service. They make it difficult to cancel the service or unsubscribe, as they “get you to lock in the purchase and not let you shop around” (P3). There are also dark patterns that restrain through hiding choice. One of the examples relates to advertising, where companies use emotion to deceive consumers: “when you are trying to opt-out from providing data for personalized advertising, you see ‘You

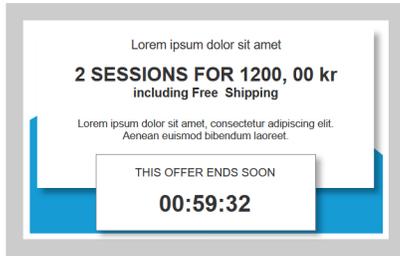


Figure 2. The example of scarcity bias-based urgency dark pattern.



Figure 3. The example of pressuring consumers to purchase through showing activity of others.

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).

4.1.2. Second-generation dark patterns. The *second-generation* patterns were more prominently discussed by our experts. The *second-generation* patterns are hard to identify by consumers and difficult to describe. Often relying on extensive data collection, they might target individual consumers and their vulnerabilities. In contrast to the *first-generation*, *second-generation* patterns are hard to notice, regulate by law, and they are less likely to self-regulate. They were perceived as more harmful because their effects are arduous to recognize.

One example of a *second-generation* dark pattern comes from the hotel industry, where a hotelier might have different types of rooms (P3). They apply one of the activity patterns (Figure 3). Consumers are encouraged to buy due to the limited number of rooms of a particular type. There is no way of knowing whether the hotelier has “a number of different room types, all of which are basically similar, and they might have segmented their rooms in a way to make it look like they are always running out of one room type or another” (P3). It is also hard to assess if a specific algorithm ensures that these room types are always prioritized and shown to the consumer first in the search results. Consequently, consumers cannot know whether they are

being manipulated, while a regulator might struggle to distinguish whether the service contains a dark pattern or not.

Privacy. Another issue discussed by our interviewees relates to privacy. The privacy-violating dark patterns (resulting in extensive data collection without the consumer knowledge, e.g., privacy Zuckering—here, a service collects more data than the consumer agreed to and shares it with other parties) are on the border between the *second-generation* and *first-generation* patterns. What brings them more towards the former is that they rely on extensive and ubiquitous data collection, which incoherent information disclosures of privacy policies or different dark patterns often prevent understanding. It is commonly acknowledged that people are often unaware of data collection and processing practices for many reasons. For instance, they might not understand how the technology works: “I think people do not understand its implications. 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).

An interesting point was raised by one of our interviewees regarding the problematics of the digital market framework based on, what they called, the “financial framework” —exchanging payment for some goods. The framework implies that if something goes wrong within the transaction, the consumer can request compensation. Considering the data collection, it might be impossible to receive data back. Regulations, e.g., General Data Protection Regulation (GDPR), guarantee consumers access to, and deletion information companies hold about them. However, it is possible that when the consumer requests their information to be deleted, it was already utilized, contributing to the company’s profits and potential consumer loss, both of which might be difficult to quantify.

Usefulness. Another issue raised by our experts considering the *second-generation* patterns was their usefulness. Some interviewees questioned whether the extensive collection of personal information and customized services are beneficial for the consumer. Personalization could have positive effects, but it also poses a challenge for consumer agencies. As P5 said: “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.” On the other hand, the extensive data might be used to exploit weaknesses, e.g., companies using dark patterns that have an effect only when specific vulnerabilities are targeted—“if you are targeting those addicted to gambling, for example, by using those big buttons containing slogans

like ‘try your luck’”(P1). Regardless, *second-generation* patterns might be perceived as useful because they contain information, which if truthful, could help to make better choices (e.g., true testimonials indicating problems with products or service)—“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). Nevertheless, when false (fake reviews hiding the product’s faults), such information might be used to prevent shopping around, reducing the probability of the optimal choice.

4.1.3. Vulnerability. Our interviewees admitted that they had little practical experience with vulnerable consumers. However, some mentioned the “traditionally” vulnerable groups, such as people with cognitive impairments, financially vulnerable, low income, specific age groups (younger and older generations), and low numeracy, might more likely become dark patterns’ victims. Still, notably in the context of *second-generation* patterns, our experts expressed concerns about the context-dependent vulnerability, which becomes an issue when companies collect extensive amounts of information about an 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, [...] and they could target you with marketing or offers that are directly linked to where you are at the moment, e.g., after divorce” (P1). The context-dependent vulnerability creates an opportunity for unethical business practices, and it enables companies to, instead of targeting one particular group of people, potentially target all consumers. The many contextual factors, such as time pressure, psychological aspects (e.g., mood, 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).

4.2. Reducing effects of dark patterns

Through the analysis of responses to interview question (b), we attempted to answer RQ2 —identify how to reduce consumer detriment and add to the market well-functioning. Here, we applied the BCW (Table 1) as a framework.

Policy categories. Policy-making—both legislation and regulation—has been the most frequently mentioned solution to overcome the effects of dark patterns. Additionally, guidelines and service provisions were

discussed by experts.

Regulation & Legislation. Our experts addressed the lack of rules about auditing and sharing the audit logs. The design process documentation could be one of the measures to prevent manipulative designs. “[An] audit in which [companies] go about the decision making for things like how they write their T&Cs, and the steps they have gone through” (P3). If applied, such audit log requirements could improve transparency and indicate whether companies have done everything to ensure that their designs do not negatively affect consumers.

Another way of reducing detriment is through regulations informed by research. Although research about dark patterns is growing, it is rarely applied as a backbone for new regulations and legislation. According to our experts, they should be based on empirical evidence, e.g., “if you have evidence of harm and market failure, I think things need to be banned” (P4). There should also be more research dedicated to the quantification of the harmful effects of dark patterns.

The outcome-based strategies to regulate the market are another way to ensure its well-functioning. In the traditional approach, regulators require companies to disclose certain information, e.g., “it must be clear, it must be placed in a relevant part of the product page [...] then if the company complies [...] everything is fine” (P5). Our interviewees criticized such an approach and proposed an improved version of it, where companies must convey certain information without receiving instructions on how this should be done. Here, companies only need to make the information displays effective. This approach would impact how the regulators work. Instead of asking companies what they did, they could focus on the actual outcomes—checking whether consumers noticed and understood the displayed information. As P5 explained: “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.”

On the other hand, such an approach could challenge the current practices that regulators apply to surveil companies, e.g., in some countries, it is not easy to purchase without a real identity, while arranging a fake identity is illegal. However, these obstacles could be subdued, “You are doing the test for an average representative sample of consumers and seeing whether firms are meeting the standard” (P8). Here, the consumers’ panels could provide insights into reading, understanding, and interacting with specific designs.

The dialogue between companies and regulators was also mentioned as a way to subdue dark patterns. Such dialogue might be a good start in identifying where the border between nudging and manipulating is—“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).

Lastly, there is a need for a precise definition of prohibited terms. Such prohibition relates mainly to T&Cs and the concept of informed consent, which might be affected by dark patterns. Among interviewees, the concept of consent was also questioned: “I just think that there is a need for a movement away from this idea of consent [...] I think, for certain things it makes sense, but for others, it does not” (P4).

Guidelines. Other theme identified in the data relates to guidelines —lack of documentation checklists. The companies could benefit from the concise information on how to document the design process. Such checklists could help to “prove that [companies] are honest” (P3). In some ways, related to the lack of checklists is the lack of precise specifications. There are no comprehensible lists of disclosure specifications: “we have to specify what information the consumer needs, and then, whenever we do that, we also have to specify how” (P5).

Guides could help especially smaller companies. Without the appropriate resources, they might not be able to assess maliciousness of their designs and instead follow any commonly used designs. Guides, templates, or design patterns could help such 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. [...]” (P8) should be applied together with outcome-based regulation.

Service provision. The last policy category surfaced from the interviews relates to service provision, particularly the lack of enforcement tools. Unlike most of the above-described solutions, this one does not directly relate to the companies but regulators. According to our interviewees, not all of the dark patterns can be analyzed with tools currently in use, which possibly 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). Such improved tools could target the texts of T&Cs to make lawyers’ work more effective or help identify misleading text-based and visual designs.

Intervention functions. The second part of the BCW relates to intervention functions (Table 1). The experts mentioned only a few intervention functions: education, coercion, modeling, and training.

Education. One way to prevent dark patterns is through a code of ethics implemented into teaching. Such code could inform designers or other people who play an active role in technology development, potentially reducing intentional and unintentional dark designs. However, education is not the soundest way of making a

change: “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 no meaningful punishment for the use of dark patterns. Our experts suggested that one way of implementing behavioral change would be by punishment or costs enforced on companies, e.g., big fines for not complying with rules.

Modeling. The results of 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 some really public cases of enforcement, of the worst offenders, would help the industry as a whole” (P4). The modeling intervention might be essential for smaller companies without a team of lawyers, which base their design-related decisions on competitors’ designs reported in news outlets.

Training. Training is the only intervention function applicable to regulators. There seems to be insufficient training in the enforcement agencies, particularly in countries with relatively well-developed consumer rights legislation, such as the EU members. For instance, enforcement agencies might lack personnel with a good understanding of the digital market and technology. Such personnel requires training to make their work more efficient and robust.

Changes to consumers’ behavior. The interviewees did not extensively discuss direct changes to consumers’ behavior. Perhaps this is because consumers should not be overloaded with information and learning. Although in some instances friction could be added to the design to activate people’s reflective thinking regarding the decision (e.g., in context of privacy-related dark pattern) [29, 30], it is not clear how designs that slow down the interaction would affect usability or experience in a different contexts (e.g., at the last stage of purchasing process). Regardless, our interviewees shifted responsibility for informed decision-making away from the consumer towards policymakers, regulators, and service providers.

Nonetheless, on few occasions, the experts mentioned that the consumers’ education and awareness could help prevent the detrimental effects of dark patterns—“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 challenging to detect. However, communication and marketing as an intervention

function might be one way to improve consumers' decisions: "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. Discussion

5.1. Harmfulness of dark patterns

Our first objective was to categorize dark patterns according to their harmfulness (RQ1). However, our interviewees, instead of focusing on the individual dark patterns, assessed their harmfulness by dividing them into two new categories—*first-generation* and *second-generation*. To some extent, *first-generation* patterns are more "traditional." Here, interviewees mentioned some of the well-known examples, e.g., hidden cost, scarcity-based patterns. It is unlikely that companies will overuse such patterns as the consumer might get annoyed and stop using the service. Harmfulness seems to be associated with how easy it is to identify dark patterns. Considering the COM-B framework and new categorization, it appears that *first-generation* patterns seem to affect consumers' motivation or capability primarily. They often lead to economic loss, restrains, and are heavily rooted in heuristics and cognitive biases. While they are harmful, they might be easier to regulate and intervene with.

The *second-generation* dark patterns might be more harmful because of their insidious nature. These patterns could be personalized, targeting specific vulnerabilities. They are based on extensive data collection, and consumers are often unaware that sensitive data is used for personalization. Additionally, in *second-generation* patterns, it is challenging to draw a border between their useful and malicious nature. They seem to affect all COM-B framework's components, including comprehension. They prey on cognitive biases and combine various malicious tactics in one mechanism that is hard to detect. The *second-generation* patterns' effects are difficult to measure, unlike, e.g., quantifiable economic loss. Considering all the above, *second-generation* patterns might be difficult to regulate or prohibit.

Our results show that dark patterns affect consumer empowerment by adding to the personal detriment. They prevent consumers from making "informed" decisions by steering their opportunities, e.g., *first-generation* patterns restricting ability to choose. Moreover, the close relationship between the *second-generation* dark patterns and privacy, which, when not protected, enables extensive data collection directed at individuals'

weaknesses, appears a particular threat to consumers. Here, consumers' empowerment is non-existent because the information needed to make optimal decisions (e.g., about potential consequences of decision) is beyond consumers' reach—hidden in companies' policies and T&Cs, or presented in a cognitively inaccessible way. Therefore, both *first-* and *second-generation* dark patterns impact digital market well-functioning, particularly affecting its social (e.g., financial loss) and technological (e.g., inaccessibility) dimensions.

Autonomous choice. Consumers have a right to make autonomous choices. Dark patterns stand in the way of autonomy, raising ethical concerns related to manipulation. Online manipulation, defined as "the use of information technology to covertly influence another person's decision-making" poses a threat in many contexts, including the commercial one, and due to its hidden nature, it requires extra vigilance [13]. Nevertheless, not all dark patterns manipulate—it appears that perhaps the *first-generation* patterns might be more likely coercive. For instance, the dark pattern called forced enrollment pushes the consumer to create an account and share their information to preview a website. Here, coercion eliminates alternative choices (e.g., previewing website without sharing data), affecting market well-being through driving an unfair and imbalanced competition (i.e., consumer locked in one service).

We believe that some of the dark patterns manipulate—if we agree that "feeling manipulated" means not understanding why we behaved the way we did or whether our actions were beneficial for ourselves or others [13]. Moreover, manipulators (e.g., online companies or designers) "control [their] victim by adjusting her psychological levers" [31]. Therefore, most dark patterns should be deemed manipulative because they frequently build on psychological effects, influence consumers' beliefs, desires, or emotions.

5.2. Overcoming barriers

We aimed to gain knowledge about the potential ways to overcome consumer detriment—empowering the consumer and ensuring fair market competition (RQ2). Using the behavioral change wheel, we classified possible solutions according to policy categories and intervention functions. The experts extensively discussed policy categories, especially regulation and legislation, providing ideas for potential improvements of the digital market. Although we used the BCW as a framework for the data analysis, identified solutions apply primarily to companies and regulators, not consumers directly. However, they carry an indirect impact on consumers'

behavior. The shift in the responsibility related to the consumer choices seems appropriate, as some research shows that end-users might fall for the manipulative designs, even when they are aware of their existence [32].

Nevertheless, regulation and legislation—the most discussed policy categories—have some disadvantages. For instance, the transparency of the design process might increase the market imbalance. Here, larger companies might have an advantage over smaller companies without the resources to comply with regulations. They may not be able to perform the testing needed to estimate whether the design has a detrimental effect on consumers. However, such limitations could be subdued by ensuring that the smaller companies receive adequate help (e.g., through the service provision) or customizing requirements depending on the company size (e.g., a longer time to fulfill the requirements).

Our findings to some extent align with the framework from [23] (published after our analysis). Here, to assess the “dark” side of designs, the normative lens should be applied, measuring individual and collective welfare, regulatory objectives, and individual autonomy. Notwithstanding a case study presented in [23], such an approach’s implementation and effectiveness remains unknown and should be evaluated. Still, considering some similarities between our experts’ opinions and the proposed framework, perhaps it is one way to assess dark patterns and apply proposed solutions in regulatory enforcement.

To sum up, dark patterns adversely affect consumers’ empowerment, leading to loss of autonomy, uninformed decisions, and potential harms (e.g., economic loss). They violate the principles of fair market competition, preventing consumers from “shopping around,” where companies utilizing dark patterns have “locked-in” consumers in their services to maximize revenues [33]. However, dark pattern-like designs might be useful if conveying true information (e.g., providing information about the room types on hotel booking websites). Therefore, the researchers’ and regulators’ challenge is to assess accurately when the designs become “dark.” This blurred borders between harmful and valuable effects of patterns require further investigation. Some regulations already target malicious designs (European Directive 2005/29/EC on unfair commercial practices (UCPD) [34] or California Consumer Privacy Act (CCPA) that directly targets dark patterns [35]) they have their limitations. For instance, the UCPD lists only some designs that could be dark patterns, and the sanctions are still left to be defined by the EU members independently and might be too small. Another issue regarding legal regulations is the problematic enforcement, for instance, how the digital market should be surveilled or how to

identify and classify dark patterns designs. Here, the more extensive research on dark patterns, especially the *second-generation*, and the development of new technologies could help.

5.3. Limitations and future work

We collected data from eight experts who expressed their subjective opinions. We combined experts from different fields and geographic areas to counterbalance this limitation, but the sample size remained small. Future research could apply different qualitative data collection methods, increasing the sample size and variability, e.g., through focus groups or workshops containing groups of experts from different scientific fields and nationalities. Such methods could assure that individual dark patterns are presented to participants, and their harmful effects are easier to assess. Additionally, future research could apply BCW to make experts familiar with the framework and then “brainstorm” their solutions accordingly.

6. Conclusion

The current research results present a new categorization of dark patterns concerning their harmful effects. The first category is the *first-generation* patterns that are harmful but relatively easy to recognize and regulate. However, another category—the *second-generation* patterns might be even more harmful because they are built on extensive data collected from unaware consumers and can target individuals’ temporary vulnerabilities. Researchers and policymakers should give this new category of dark patterns more attention. Further, our findings denote possible guidelines to tackle the adverse effects of dark patterns on consumers and sustain a well-functioning digital economy. In particular, we explain how the four policy categories (regulation, legislation, guidelines, and service provision) and the four intervention functions (education, coercion, modeling, training), which, if implemented, could change the behavior of regulators and online companies, consequently improving consumers’ well-being.

Acknowledgment

This study was a part of the project financed by Konsumentverket (Swedish Consumer Agency).

References

- [1] United Nations, “Goals.” <https://sdgs.un.org/goals>. accessed: 05-2021.
- [2] S. Human, G. Neumann, and R. Alt, “Human-centricity in a sustainable digital economy,” 2021.

- [3] P. Gazzola, G. Colombo, R. Pezzetti, and L. Nicolescu, "Consumer empowerment in the digital economy: Availing sustainable purchasing decisions," *Sustainability (Switzerland)*, vol. 9, no. 5, 2017.
- [4] European Commission, "Consumer Empowerment in the EU," no. July, p. 15, 2011.
- [5] European Commission, "Study on measuring consumer detriment in the European Union Final report Part 1- Main report," tech. rep., 2017.
- [6] J. Luguri and L. Strahilevitz, "Shining a Light on Dark Patterns," *SSRN Electronic Journal*, pp. 1–51, 2019.
- [7] 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.
- [8] C. Cara, "Dark patterns in the media: A systematic review," *Network Intelligence Studies*, vol. 7, no. 14, pp. 105–113, 2019.
- [9] 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, pp. 1–14, 2018.
- [10] A. Mathur, G. Acar, M. J. Friedman, E. Lucherini, J. Mayer, M. Chetty, and A. Narayanan, "Dark patterns at scale: Findings from a crawl of 11K shopping websites," *Proceedings of the ACM on Human-Computer Interaction*, vol. 3, no. CSCW, 2019.
- [11] D. M. Oppenheimer and E. Kelso, "Information processing as a paradigm for decision making.," *Annual review of psychology*, vol. 66, no. 1, pp. 277–94, 2015.
- [12] OECD, "Improving Online Disclosures With Behavioural Insights," *OECD Digital Economy Papers*, no. 269, 2018.
- [13] D. Susser, B. Roessler, and H. Nissenbaum, "Online manipulation: Hidden influences in a digital world," 2018.
- [14] H. Brignull, "Dark patterns." <https://www.darkpatterns.org/>. accessed: 05-2021.
- [15] 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.
- [16] D. Kahneman, "A Perspective on Judgment and Choice," *American Psychologist*, vol. 3, no. 4, pp. 7–18, 2003.
- [17] J. S. B. 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.
- [18] R. Thaler and C. Sunstein, *Nudge. Improving Decisions About Health, Wealth, and Happiness*. 2008.
- [19] G. Gigerenzer and W. Gaissmaier, "Heuristic Decision Making," *Annual Review of Psychology*, vol. 62, no. 1, pp. 451–482, 2011.
- [20] N. H. Lurie and C. H. Mason, "Visual representation: Implications for decision making," *Journal of marketing*, vol. 71, no. 1, pp. 160–177, 2007.
- [21] L. Deng and M. S. Poole, "Affect in web interfaces: A study of the impacts of web page visual complexity and order," *MIS Quarterly: Management Information Systems*, vol. 34, no. 4, pp. 711–730, 2010.
- [22] 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.
- [23] A. Mathur, M. Kshirsagar, and J. Mayer, "What makes a dark pattern... dark? design attributes, normative considerations, and measurement methods," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–18, 2021.
- [24] S. Michie, L. Atkins, R. West, *et al.*, "The behaviour change wheel," *A guide to designing interventions. 1st ed. Great Britain: Silverback Publishing*, pp. 1003–1010, 2014.
- [25] T. Wastling, F. Charnley, and M. Moreno, "Design for Circular Behaviour: Considering Users in a Circular Economy," *Sustainability (Switzerland)*, vol. 10, no. 1743, 2018.
- [26] Konsumentverket, "Underlagsrapport 2021:1 Barriers to a well-functioning digital market," 2021. <https://www.konsumentverket.se/globalassets/publikationer/produkter-och-tjanster/ovriga-omraden/underlagsrapport-2021-1-barriers-digital-market-konsumentverket.pdf>.
- [27] S. Döringer, "'The problem-centred expert interview'. Combining qualitative interviewing approaches for investigating implicit expert knowledge," *International Journal of Social Research Methodology*, vol. 24, no. 3, pp. 265–278, 2021.
- [28] A. Bogner, B. Littig, and W. Menz, "Generating qualitative data with experts and elites," *The SAGE handbook of qualitative data collection*, pp. 652–667, 2018.
- [29] A. Terpstra, A. P. Schouten, A. de Rooij, and R. E. Leenes, "Improving privacy choice through design: How designing for reflection could support privacy self-management," *First Monday*, 2019.
- [30] A. Terpstra, "Think before you click: how reflective patterns contribute to privacy (position paper)," in *What Can CHI Do About Dark Patterns? CHI Workshop*, 2021.
- [31] R. Noggle, "Manipulative actions: a conceptual and moral analysis," *American Philosophical Quarterly*, vol. 33, no. 1, pp. 43–55, 1996.
- [32] K. Bongard-Blanchy, A. Rossi, S. Rivas, S. Doublet, V. Koenig, and G. Lenzini, "I am definitely manipulated, even when i am aware of it. it s ridiculous!–dark patterns from the end-user perspective," *arXiv preprint arXiv:2104.12653*, 2021.
- [33] W. Browne and M. S. Jones, "What works in e-commerce-a meta-analysis of 6700 online experiments," *Qubit Digital Ltd*, vol. 21, 2017.
- [34] Council of European Union, "Directive 2005/29/ec of the european parliament and of the council," 2005. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32005L0029&from=EN>.
- [35] Attorney General Rob Bonta - Press Release, "Attorney General Becerra Announces Approval of Additional Regulations That Empower Data Privacy Under the California Consumer Privacy Act," 2021. <https://oag.ca.gov/news/press-releases/attorney-general-becerra-announces-approval-additional-regulations-empower-data>.