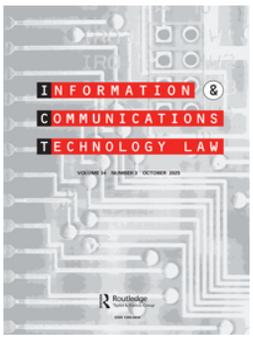

Routledge
Taylor & Francis Group

Information & Communications Technology Law

ISSN: 1360-0834 (Print) 1469-8404 (Online) Journal homepage: www.tandfonline.com/journals/cict20

No harm no foul: how harms caused by dark patterns are conceptualised and tackled under EU data protection, consumer and competition laws

Cristiana Santos, Viktorija Morozovaite & Silvia De Conca

To cite this article: Cristiana Santos, Viktorija Morozovaite & Silvia De Conca (2025) No harm no foul: how harms caused by dark patterns are conceptualised and tackled under EU data protection, consumer and competition laws, Information & Communications Technology Law, 34:3, 329-375, DOI: 10.1080/13600834.2025.2461958

To link to this article: https://doi.org/10.1080/13600834.2025.2461958

Published online: 17 Feb 2025.

INFORMATION & COMMUNICATIONS TECHNOLOGY LAW
2025, VOL. 34, NO. 3, 329–375
https://doi.org/10.1080/13600834.2025.2461958

Routledge
Taylor & Francis Group

No harm no foul: how harms caused by dark patterns are conceptualised and tackled under EU data protection, consumer and competition laws

Cristiana Santos[a], Viktorija Morozovaite[b] and Silvia De Conca[c]

[a]Law and Technology, Department of International and European Law School of Law, Utrecht University, Utrecht, Netherlands; [b]EU Law, Amsterdam Centre for European Law and Governance (ACELG), University of Amsterdam, Amsterdam, Netherlands; [c]Law & Technology Transnational Legal Studies Department, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

ABSTRACT

Although several Human–Computer Interaction (HCI) studies have empirically investigated the harms caused by dark patterns, with policymakers and regulators regarding these harms significant, they have yet to be examined from a legal perspective. This paper identifies the individual, collective, material and non-material harms deriving from dark patterns, dissecting the role that harms play in the emerging European 'dark patterns acquis', comprising the Digital Services Act, Digital Markets Act, AI Act and Data Act. In particular, it systematises the body of knowledge of dark patterns' harms from HCI scholarship and proposes a dark pattern harm taxonomy. Ultimately, the paper reconciled the debate concerning dark patterns' harms in HCI with the legal requirements for assessing harms, in light of the remedies mechanisms offered by European data protection, consumer law and competition law.

KEYWORDS

Dark patterns; harm; redress; GDPR; UCPD; competition; EU Laws

1. Introduction

Dark patterns[1] refer to online user interface 'practices that materially distort or impair, either on purpose or in effect, the ability [of users] to make autonomous and informed choices or decisions.'[2] Such design choices benefit the online platform using them, but might be contrary to the interest of a user. Common examples are the cookie pop-ups of websites displaying a prominent, bright 'accept' button and hiding the 'reject' one, or e-commerce platforms pressuring visitors into buying with unsupported claims of scarcity ('only one item left!'). Dark patterns are used by designers and developers of digital products or services to increase sales, collect data, or simply grab user attention.

The magnitude of this phenomenon is significant and likely to grow in digital markets: an investigation by the EU Commission and the consumer protection authorities of

CONTACT Cristiana Santos c.teixeirasantos@uu.nl

[1]We interchangeably use the terms 'dark patterns' and 'deceptive design'.

[2]Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) [2022] OJ L 277/1, recital 67.

23 Member States found that over 40% of analysed e-commerce websites deployed it,[3] while research is beginning to shed light on its diffusion in Internet of Things devices too.[4] Due to the permeation of information technologies, manipulative practices can be implemented at low costs, large scale and with unprecedented sophistication in interactive, intrusive, and adaptive environments, which increases their efficacy.

To investigate the impact of dark patterns, empirical research has been evaluating both users' perceptions and harms, and their presence across digital services and modalities, including mobile apps, websites, voice assistants, e-commerce sites, social networks, games, and privacy control mechanisms like consent banners.[5] There is a growing body of empirical evidence confirming that dark patterns elicit or lead to actual or potential harm (e.g. privacy harm, autonomy loss, intensive labour, cognitive effort, time consumption, attentional harm, emotional distress, mental harm, financial loss, and collective harms),[6] with several new types of dark patterns' harms being recently acknowledged in scholarly works (e.g. emotional load, social injustice).[7]

On the regulatory side, dark pattern practices have been subject to administrative and judicial interventions,[8] and are explicitly mentioned by some regulators in their decisions.[9] However, to date, the harms caused by dark patterns have not been addressed in national case law[10] and users are still exposed to these practices with little remedy. In the last four years, dark patterns have been the focus of EU regulation and policymaking. Initially, the regulatory interventions against dark patterns orbited around consumer protection and personal data protection. In 2022, various regulatory bodies, including the US Federal Trade Commission (FTC), the UK Competition and Market Authority (CMA), the Dutch Consumer and Competition Authority (ACM), the European Commission, and other stakeholders, such as the European Data Protection Board (EDPB), the Organisation for Economic Co-operation and Development (OECD), and the European Consumer Organization (BEUC) have issued high-profile policy guidance on dark patterns and deceptive design, acknowledging that dark patterns trigger a unique potential for impact on user behaviours, including undermining of user agency and disparate impacts on vulnerable or disempowered communities.[11] Some regulators have also acknowledged the collective dimension of harms to well-functioning and competitive markets.

[3]European Commission, 'Consumer protection: manipulative online practices found on 148 out of 399 online shops screened' (Press Release, 30 January 2023) <https://ec.europa.eu/commission/presscorner/detail/en/ip_23_418> accessed 14 June 2024.

[4]See the IoT study listed in Annex A under [ST-12].

[5]J Gunawan, A Pradeep, D Choffnes and others, 'A Comparative Study of Dark Patterns Across Web and Mobile Modalities' (2021) 5 *Proceedings ACM Human-Computer Interaction* CSCW2 Article 377.

[6]J Gunawan, D Choffnes, W Hartzog and C Wilson, 'Towards an Understanding of Dark Pattern Privacy Harms' (2021) *Proceedings of the ACM Human Computer Interaction*, CHI21, May 8–13, 2021, Online Virtual Conference. See also the OECD report listed in Annex A under [PR-0].

[7]I Chordia, LP Tran, TJ Tayebi and others, 'Deceptive Design Patterns in Safety Technologies: A Case Study of the Citizen App' (2023) *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (CHI 23), ACM.

[8]M Leiser, C Santos and K Doshi, 'Legal Cases' (Deceptive Patterns, 2023) <https://deceptive.design> accessed 10 September 2024.

[9]C Santos and A Rossi, 'The Emergence of Dark Patterns as a Legal Concept in Case Law' (2023) *Internet Policy Review* 1; M Leiser and C Santos, 'Dark Patterns, Enforcement, and the Emerging Digital Design Acquis: Manipulation beneath the Interface' (2024) 15(1) *European Journal of Law and Technology*.

[10]J Gunawan, C Santos and I Kamara, 'Redress for Dark Patterns Privacy Harms? A Case Study on Consent Interactions' (2022) *Proceedings of the 2022 Symposium on Computer Science and Law* (CSLAW '22), ACM, 181, 188.

[11]CM Gray, CT Santos, N Bielova and T Mildner, 'An Ontology of Dark Patterns Knowledge: Foundations, Definitions, and a Pathway for Shared Knowledge-Building' (2024) *Proceedings of the CHI Conference on Human Factors in Computing Systems* (CHI '24), ACM, Article 289, 1.

In 2023, the term 'dark patterns' was codified into EU law, notably by the new Digital Services Act (DSA), introducing provisions aimed at safeguarding users against deceptive or manipulative user interfaces.[12] The list of provisions dealing with dark patterns, directly or indirectly, is destined to grow, with the Digital Markets Act[13] (DMA) and the recent Artificial Intelligence Act[14] (AI Act) and the Data Act[15] (DA) addressing, respectively, manipulative AI systems and deceptive design used in the context of data sharing. There is thus a growing consensus that, in digital environments, choice architecture plays an important role in facilitating legal compliance. However, as digitalisation intersperses every aspect of modern lives, there is also a growing need for cohesion in the increasing, yet fragmented, regulatory responses to dark patterns' harms.

Whilst other disciplines debate the nature and incidence of various types of harms, there is the necessity, from the legal perspective: (i) for dark patterns harms to be reconciled with the legal approach to harm or damage, wherein a causal link between infringement and damage is required, and where commonly recognised losses are mostly patrimonial and individual;[16] and (ii) for an alignment of dark patterns harms within each respective legal field. To fill this gap in literature, this article aims to answer: *to what extent are dark patterns' harms*[17] *legally recognisable and enforceable in EU law?*[18]

For the purposes of this article, our legal analysis of harms focuses on data protection law, consumer law and competition law. We chose these legal fields because when it comes to dark patterns' harms, they fall within the competences of the respective national authorities. In fact, both data protection and consumer protection authorities have already sanctioned harmful dark pattern practices in these domains. With growing concerns over the negative impact of dark patterns on competitive markets, competition law provides a logical legal avenue to deal with them. In addition, all the respective regulators in these domains have already issued policy reports regarding dark patterns. Nevertheless, to date, harm as a constitutive element of dark patterns has been overlooked. Finally, while the main focus of our legal analysis revolves around data protection law, consumer protection law and competition law, we nevertheless recognise the

[12]Digital Services Act (n 2), Article 25.

[13]Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), OJ L 265, 12.10.2022, p. 1–66.

[14]Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

[15]Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act).

[16]Gunawan et al (n 10).

[17]For the purposes of this study, harm should be intended as the impairing of or detriment to a right or (lawful) interest of an individual or group. Damage is to be intended as any type of detriment to an individual's sphere; loss is a type of detriment, consisting of economic loss (loss of income, costs incurred, etc) or non-economic loss (pain, suffering, negative effects on quality of life). See C von Bar, E Clive, H Schulte-Nölke and others, 'Principles, Definitions and Model Rules of European Private Law. Draw Common Frame of Reference' (2008) Interim Edition, para 64 https://www.ccbe.eu/fileadmin/speciality_distribution/public/documents/EUROPEAN_PRIVATE_LAW/EN_EPL_20100107_Principles__definitions_and_model_rules_of_European_private_law_-_Draft_Common_Frame_of_Reference__DCFR_.pdf accessed 10 September 2024.

[18]Note that the legal analysis of this article is focused on how harms are defined and treated in the respective legal domains. Thus, our application of the EU laws in relation to dark patterns' harms does not include the normative propositions of how dark patterns are defined or treated in those laws.

importance of the emergent EU digital regulatory frameworks dealing with dark patterns: DSA, DMA, AI Act and Data Act. We thus analyse the role of harm in these legal areas too.

We make the following contributions in this paper:

(1) We provide a fine grained taxonomy of dark patterns' harms based on an in-depth analysis of 39 sources (12 policy reports and 27 academic studies), and propose a three-level taxonomy across low-, meso-, and high-level harms, including individual and collective, material and non-material harms, amounting to a total of 20 types of harm;
(2) We appraise these harms in light of the legal requirements of various EU legal regimes – data protection, consumer law, competition law. In doing so, we describe the state of the art when it comes to different legal regimes' treatment of harm vis-a-vis dark patterns' harms. By conceptualising harms, we provide the foundational steps for measuring intangible and material harms caused by dark patterns[19] which could help enforcement authorities define their priorities;
(3) We dissect the role that harms play in the emerging European 'dark patterns acquis', under the Digital Services Act, Digital Markets Act, AI Act and Data Act and conclude that most non-material harms, while commonly discussed in empirical and policy studies, are often challenging to redress and difficult to substantiate;
(4) We suggest recommendations for regulators, policymakers and social scientists based on our findings.

2. Methodology

To organise the harms in a taxonomy,[20] we drew upon a total of thirty-nine sources, comprising twelve policy reports and twenty-seven academic literature studies. The sources analysed are listed in Annex A, divided into two tables, one for policy reports and one for academic studies. Each source is associated with an identifier composed of an acronym (PR for policy reports and ST for academic studies) and a number. To maintain this article readable, throughout its text the sources are indicated in square brackets with their identifier. The complete list can be consulted in Annex A at the end of the article. Additionally, Annex C offers a visual representation of the taxonomy.

Few taxonomies of dark pattern harms are found in the existing scholarship. Mathur et al. [ST-19] provides a first brief general classification of end-user harms in their description of individual welfare-based normative perspectives of dark patterns. This includes Financial Loss, Invasion of Privacy, and Cognitive Burden, within an individual welfare lens; from a collective welfare perspective, the authors identify Competition, Price Transparency and Trust in the Market. In contrast to the welfare perspective, the authors add Individual

[19]F Esposito and A-L Sibony, 'In Search of the Theory of Harm in EU Consumer Law: Lessons from the Consumer Fitness Check' in Klaus Mathis and Avishalom Tor (eds) *Consumer Law and Economics* (Springer 2020), 5.

[20]Our paper provides a taxonomy of harms rather than a typology because it aims to systematically classify and organise the different types of harms within a hierarchical structure based on their common features, relationships with other types of harms, either at low, meso or high-level of harms, and considering their intrinsic characteristics. The proposed taxonomy allows for a comprehensive and detailed categorization, which helps in understanding the various dimensions of the harms more effectively. This approach goes beyond simply listing types (as in a typology) and instead offers a structured framework that can be used for deeper analysis and application in various contexts.

Autonomy within a regulatory and rights-based lens. Gunawan et al.[21] further refined each of these general categories based on state-of-the-art studies of harms. Notably, we observed the available taxonomies used different categorizations, based on the sector they sought to describe or support. This is the case regarding the CMA and ICO taxonomy [PR-10], mostly focused on consumers and competition, while the OECD [PR-0] provided its own list of harms, which is a consolidation effort based on former works, though it excludes certain harms, such as Addictive Design, Cognitive Load, or Resignation.

Through this comprehensive analysis, we identified three hierarchical levels of harm associated with dark patterns. This effort supports stakeholders and also scholars with a shared and consistent vocabulary, which they can use to discuss harms across domains, contexts and to shape both future research, regulatory action, and redress at national level. The following sections describe our iterative steps.

2.1. Collection of scholarly empirical literature

We collected scholarly works using the snowballing method, a customary practice for legal empirical studies. We selected articles including the terms *dark patterns*,[22] *manipulative* or *deceptive design*, and *harms* either in their titles, abstracts, or keywords. To best represent the harms caused by dark patterns, we focused on international studies, in English. We elicited both empirical and non-empirical scholarly works.[23] Regarding *empirical* works, we selected Human Computer Interaction (HCI) studies (listed in Annex A) based on thematic content analysis, or user studies methods providing evidence of dark patterns' harms. Concerning *non-empirical literature*, we selected studies on dark patterns by Mathur et al. (2021) [ST-19] and Roffarello et al. (2023) [ST-22], due to their foundational, state-of-the-art contributions. We collected policy reports from stakeholders and regulators in the EU, UK, Australia and USA referring to dark patterns' harms from European Commission [PR-8], OECD [PR-0], ICO [PR-10], FTC [PR-9] EDPB [PR-7], CMA [PR-5], Norwegian Consumer Council (NCC) [PR-11], Australian Competition and Consumer Commission (ACCC) [PR-1], Consumer Policy Research Centre (CPRC) [PR-6], ACM [PR-3], and the European Parliament [PR-4]. When a type of harm and sector is sufficiently represented by recent literature, and/or saturation is reached (e.g. competition-based reports from the CMA or ACM), we limit our analysis to the most recent works and exclude older ones.[24] These academic and regulatory studies provide a strong foundation for our taxonomy of harms.

2.2. Classification of non-legal harms from empirical works

We used the following iteration, mostly inspired by Gray et al. (2024),[25] to identify existing taxonomy components, their source, and similarities between components across typologies.

[21]Gunawan et al (n 10).
[22]In this paper we use the general understanding of the definition of dark patterns and exclude broader definitions that are related to misinformation or deep fakes.
[23]We did not distinguish between empirical and non-empirical studies when examining collective harms.
[24]For example, we excluded from our analysis older research e.g., The Office of Fair Trading, 'OFT Report: Online targeting of advertising and prices' (2010), https://www.scl.org/1839-oft-report-online-targeting-of-advertising-and-prices/, accessed 23 September 2024.
[25]Gray et al (n 11).

2.2.1. Clustering similar types of harms

We listed all the harms mentioned or discussed in the aforementioned sources. Subsequently, we grouped harms that appeared either to be identical or similar (in a is-a or equivalent-to relationship) using the definitions provided by the sources themselves to identify affinities in the harms that did not have identical names. The grouping was done prevalently on excel and, for a more visual representation, we took the different groups of identical harms and created boards on Miro.[26] This visual overview allowed us to better identify recurring elements and differences in terminology within each group, as well as to highlight conceptual connections among groups of harms. It also showed that some harms appeared more often than others. For example: autonomy loss, privacy loss, loss of trust, addiction, loss of time, financial loss, and distortion of competition were some of the largest groups.

2.2.2. Creating meso- and low-level harms

Meso-level harms were most likely to co-occur across multiple sources. For instance, Psychological Detriment is present in many of the policy reports consulted. While most sources would mention adverse psychological effects without further specification, some sources described specific ways in which Psychological Detriment could manifest, such as emotional distress in the form of embarrassment, guilt, or worry [ST-3; PR-10], or even an addiction [PR-4; ST-21]. Therefore, emotional distress and addiction represent low-level harms that can be traced back to meso-level harms. Low-level harms indicate specific harms that are derived from meso-level harms and as such can be considered through the lens of a child in a parent–child relationship.[27] At this stage, we also noticed that some of the harms discussed in the literature and included in the first mapping were mentioned by very few sources in rather generic terms, or were hypothetical harms mentioned by the individuals participating in the studies, with little to no substance to them. This happened in the case of Discrimination, mentioned only by the FTC as the result of price personalisation, and Eco-chambers/Bubbles. The latter was mentioned in one source by the subjects of the study itself, who were most likely conflating different issues with dark patterns. For these reasons, we decided to exclude these harms from the final taxonomy. Our taxonomy portrays 11 meso-level harms and 5 low-level harms, which are discussed below (see also Annex C).

2.2.3. Adding the high-level harms

Finally, we noticed that the meso-level harms could be reconducted to two fundamental pairs: material and non-material (or tangible and intangible) harms, and individual and collective harms. Notably, collective harms in the literature broadly refer to 'collective welfare ' – a notion closely tied to 'consumer welfare' as understood in EU competition law and consumer protection law [PR-8; ST-9; ST-17; ST-19]. We named these more abstract types of harms high-level. The taxonomy includes four high-level harms: individual (material and non-material), and collective (material and non-material). Notably, these categories are based on a legal perspective: for example, the distinction between material and non-material harms existing in all Western legal systems in the context of civil liability

[26]Miro.com is an online tool for collaboration. It allows the creation of digital whiteboards, and it is commonly used in project management, data analysis, and creative projects.

[27]The terminology adopted to differentiate low level harms is not indicative of the severity of the harm.

(contractual/tort liability).[28] The inclusion of the high-level layer in the taxonomy is a first step into bridging the distance between harms as conceptualised by HCI experts, and harms from a legal perspective. It is also in line with the general approach emerging from the policy reports analysed.

2.2.4. Harms out of scope

In this article, we make a distinction between individual and collective harms. In the context of collective harms, we could further distinguish between market-related impacts and societal impacts. Since societal impacts may relate to content moderation decisions and harms that fall outside the scope of our article (e.g. Discrimination, Filter Bubbles), we decided to exclude them from our current analysis. We also excluded misinformation and deepfakes from the scope of our work since they refer to broader conceptualizations of manipulation and have a potential to lead to societal harms.

3. Taxonomy of dark pattern's harms

This section presents our proposed taxonomy of the harms caused by dark patterns. The taxonomy is included in Annex C and, below, information is provided, for each identified harm, on the:

(1) recurrence in the studied literature,
(2) most complete (available) definition or characterisation of a given harm according to the sources,
(3) examples of specific dark pattern(s) that trigger it, according to the sources,[29]
(4) other related harms, and
(5) associated remarks on possible difficulties and issues connected with such harm.

The section is structured into individual non-material harms (Section 3.1), individual material harms (Section 3.2), individual harms that can manifest as either material or non-material (Section 3.3), collective material harms (Section 3.4), and collective harms that can manifest as either material or non-material(Section 3.5). We further discuss the attributes of these harms (Section 3.6) and the challenges and limitations of the classification (Section 3.7).

3.1. Individual non-material harms

3.1.1. Loss of autonomy (meso-level)

Definition Ten studies referred to this harm. Loss of Autonomy refers to the user's lack of capacity to have meaningful control over their choices and make free, informed, and

[28]While this distinction existed for a long time, non-material harms (as the current terminology goes) were compensated only in very limited cases. Non-material damages refer to 'losses which do not relate to a person's assets, wealth or income and, as such, cannot be quantified in an objective manner by reference to a market price or value.' See: Case C-371/12 *Enrico Petillo and Carlo Petillo v Unipol Assicurazioni SpA* [2013] ECLI:EU:C:2013:652, Opinion of AG Wahl, para 38.

[29]Sometimes dark patterns practices are described, but not explicitly named in the studies, e.g., 'Notifications' [ST-24]. We did not subsume such practice into a dark pattern type and maintained the type of practice referred to in the study. Whenever it is very evident, we correlate a harm with specific types of dark patterns to provide a relational perspective.

meaningful choices or decisions [PR-10, p. 8–9]. From the analysed studies, most dark patterns undermine individual autonomy [PR-8]. This harm is mostly evident in privacy and also e-commerce contexts [ST-10].

Dark patterns Concerns about autonomy echo through nearly the entire dark patterns literature [ST-19, p. 13]. For example, Interface Interference limits users' agency, since it overrides users' goals with other interests, in particular, in the context of social media networks [ST-21, p. 8].[30] Techniques using Social Proof influence consumers' behaviour by accelerating the selection and purchasing process [PR-2, p. 37]. Accordingly, consumers consequently look less critically at the product, and are less inclined to compare products. Autonomy Loss is mostly evident in dark patterns that promote Addiction, such as Autoplay and Infinite Scrolling. As posited by the EC, the attention cycle that is designed to increase the amount of time spent on online platforms, can invade such decisional space [PR-8, p. 92].

Related harms As Mathur et al. posit, the vast majority of dark patterns attempt to undermine individual autonomy and a notable example are those dark patterns that enable Addiction [ST-19, p. 13]. In the sources, Loss of Autonomy is also discussed in connection with the processing of personal data, and thus related to the Loss of privacy harm, wherein: 'users may feel powerless to stop the use of their personal information in ways they do not want' [PR-10, p. 10]. On this point, the European Commission report also refers to loss of autonomy and privacy [PR-8, p. 6] and takes the view[31] that manipulation reduces user's autonomy by invading individual decision-making, and such tactics implicate so-called 'decisional privacy', comparable to autonomy, and is eroded when manipulation invades internal thought processes, reduces free will, or interferes with a user's self-interest.

Remarks Loss of Autonomy can be viewed both as an individual, but also a collective type of harm. In fact, while autonomy is defined as an individual non-material harm, if individuals feel a loss of autonomy or a lack of control over their environment, they may seek to exert control in other ways, which can lead to broader societal problems.[32]

3.1.2. Loss of trust (meso-level)

Definition Loss of Trust is not defined by any but one source, among those analysed. Only [PR-5] defined trust as 'the subjective belief of a "favourable expectation regarding other people's actions and intentions" (Waldman, 2016)'. While it is not discussed by many sources (only 7 out of 39), a significant convergence can be noted in the way this harm is discussed. In all sources, a loss of trust/diminishing of trustworthiness is reported after individuals are exposed to dark patterns.

Dark patterns While the sources analyse the impact of different types of dark patterns, there seems to be consensus over the fact that, according to the data, Loss of Trust does not manifest in the same ways for all dark patterns: it occurs with varying intensity, and in some cases it does not manifest at all. For example, [PR-8] shows that individuals self-

[30]The authors refer to types of dark patterns that fit Interface Interference strategy scope, which are (1) decision uncertainty; (2) labyrinthine navigation; (3) and redirective conditions. These dark patterns are related to the context of Social Media Networks.

[31]The EC confirms Day's view on the correlation between autonomy and privacy [ST-7].

[32]C Busch and A Fletcher, 'Harmful Online Choice Architecture' (2024) Center on Regulation in Europe (CERRE) Report, https://cerre.eu/wp-content/uploads/2024/05/CERRE-Final-Report_Harmful-Online-Choice-Architecture.pdf accessed 10 September 2024.

reported a lower level of trust after exposure to Forced Action and Personalisation (hand-in-hand with less information understanding, more frustration, more perceived aggressiveness of website, more perception of being manipulated). However, there was no Loss of Trust after exposure to Interface Interference and Confirmshaming (even though participants still recorded a lower understanding of information provided and some perception of being manipulated). Similarly, in [PR-1], some participants found the information not trustworthy after being exposed to Scarcity or Activity Claims (respectively 26 and 27% of participants), but more participants lost trust in the business after being exposed to Forced Continuity (39%) or dark patterns aiming at collecting more data (33%). Our analysis reveals that the dark patterns (among those tested by the sources) belonging to the High-level categories of Forced Action are associated with a Loss of Trust in the brand or business, for example Forced Continuity and Privacy Zuckering [PR-8]. Roach Motel, belonging to the Obstruction family was also associated with the Loss of Trust in the brand [PR-11]: this does not surprise, seeing how it shares certain effects and characteristics with the Forced Action patterns, such as the perceivable absence of choice and control, and the forcing of users into a predetermined channel of actions. Conversely, Social Engineering patterns are associated with the Loss of Trust in the information provided on a product or service, as it has been reported for Scarcity or Activity Claims [ST-11].

Related harms Overall, from the analysed sources it appears that Loss of Trust is not treated as a harm in and of itself, but as an effect of dark patterns, which is conductive to other harms, most frequently collective ones, such as loss of trust in the market and consequently harms to the economy.

In this regard, multiple sources differentiate between the Loss of Trust in the brand or business/company altogether, and the Loss of Trust in the information provided about a specific product or service.

Remarks Some contextual factors are reported as playing a role in the intensity or even in the manifesting of Loss of Trust, for example whether the deception comes from a company or a public authority [PR-5]. It is important to consider that Loss of Trust might not be treated as a potential individual harm, but as a mere side-effect of dark patterns, whose true harm resides in the risks it brings to the economy and markets. As such, it would be better addressed by competition law, or administrative sanctions, than by tort remedies.

3.2. Individual material harms

3.2.1. Physical/bodily harms (meso-level)

Physical or bodily harms are only discussed by one policy report and two experimental studies (3 sources out of 39). Notably, even in the case of experimental studies, physical or bodily harm is discussed either as a hypothetical scenario or outcome of dark patterns [ST-13], or because participants mention it spontaneously, especially as a consequence of filter bubbles [ST-3]. The 2022 Draft Report on addictive design also mentions the possibility of adverse effects of addiction on the health of individuals, but in a more indirect way, for example as a side effect of sedentary life or of excessive screen time. Overall, Physical or bodily harms deriving from dark patterns seem more like a remote possibility than a concrete harm. Should such a possibility actualise, most national tort systems have

mechanisms in place that could ensure redress, both for the economic and non-economic aspects that physical harms entail. Tackling the (remote) chance of physical harms deriving from dark patterns for precautionary or deterrence reasons seems premature, if not unnecessary. Article 5(1) of the draft AI Act, for example, mentioned physical harms explicitly, but the draft text has already been criticised as excessive and amended in subsequent iterations.[33]

3.2.2. Financial loss (meso-level)

Definition Financial Loss is among the most discussed harms in policy documents and scholarly works (14 out of 39 sources). While the majority of examined sources do not define financial loss, it falls within 'economic' or 'financial' harms related to 'individual welfare' [respectively PR-8; ST-24; ST-17]. [PR-8] provides a broad definition of financial loss, which generally relates to situations 'when dark patterns cause a financial loss as a consequence of a purchase that the consumer would not have done [otherwise]' (p. 90). Financial Loss is also closely linked to monetary dark patterns, which lead to users regretting spending money, losing track of how much they (will) spend or not knowing how much money is required to move forward [ST-16].

Dark patterns According to the literature, there are two types of sources of economic harm: dark patterns that induce consumers to purchase products that they would not have otherwise, leading to inefficient resource allocation, and dark patterns that may allow companies to charge more for their products [PR-8, p. 90]. Some of the most commonly discussed dark patterns in relation to Financial Loss include Forced Continuity, Confirm-shaming, Disguised Ads and Hidden Costs [PR-6, p. 12, 19, 21], Sneak in the Basket, Drip Pricing [PR-9], Hidden Subscription [ST-12, p. 11]. Other, less commonly discussed, ones included Currency Confusion, Monetised Rivalries and Pay to Skip [ST-16].

Related harms Financial Loss is treated as a harm by the sources. Nevertheless, it is intertwined with other harms, including Addiction [PR-4] and Discrimination [PR-9].

Remarks The impact on consumers caused by dark patterns is particularly difficult to quantify because it varies on a case by case basis [ST-8]. That is not surprising, considering that different consumer attributes such as, *inter alia*, budgetary or time constraints and vulnerability could affect their decision-making. In addition, financial loss can also occur as a result of dark patterns deployed across different modalities, which are less explored in the literature and regulatory frameworks. One of the sources exemplified this by the unique implications of dark patterns in IoT context [ST-12].

3.3. Individual harms that can manifest as either material or non-material

3.3.1. Privacy loss (meso-level)

Definition This harm is referred to by most policy studies and it is the most studied harm among the sources (15 studies out of 39). It is defined as the choices about user's personal information that do not align with their preferences, such as sharing more personal information than they would otherwise volunteer [PR-10, p. 10]. As a result of online choice architecture in a privacy context, consumers may end up sharing more personal data

[33]S De Conca, 'The Present Looks Nothing Like the Jetsons: Deceptive Design in Virtual Assistants and the Protection of the Rights of Users' (2023) 51 *Computer Law & Security Review* 1.

or allow it to be used more extensively than intended. This harm can be worsened and lead to more extensive processing about user's behaviour, preferences and attitudes and, ultimately, unwanted targeted advertising or profiling [PR-10, p. 10].

Dark patterns Our analysis of the sources revealed that Privacy Loss can be caused by several types of dark patterns practices. This harm can occur through nagging notifications in social media and video game friend spam [PR-10, p. 10], and in obstructive choices mechanisms [PR-9, p. 15]. It is most commonly identified in contexts such as cookie banners and account registrations. Concerning cookie banners, the dark patterns inducing Privacy Loss refer to visual manipulation [PR-5, p. 85], and default options that undermine consumers privacy [ST-19]. Another study mentioned that manipulations of cookie banners increase consent by 17% of the sample [ST-1]. In the context of subscription accounts, Privacy Loss is reported when users are required to create accounts to use or set up a device (known as forced registration) and when it is difficult or impossible to delete an account [ST-12, p. 19].

Related harms Privacy Loss was mainly related to Loss of Autonomy, as it prevents users from effectively managing their privacy both while using a device and after they stop using it [ST-12, p. 19].

Remarks According to the analysed studies, Privacy Loss is discussed from both an individual and collective perspective, affecting all the participants of a given service/platform whose data was processed and analysed in a concrete setting.

3.3.2. Psychological harm (meso-level)

This meso-level category consists of four low-level harms: Emotional Distress, Addiction, Labour and Cognitive Burden, Attentional Harms.

3.3.2.1. Emotional distress (low-level). **Definition** This harm is widely recognised, being mentioned in 15 sources. Emotional Distress is mostly associated with annoyance and frustration, but strongly connected to resignation and disengagement. Related emotional impacts can refer to pressure, mood change [ST-17, p. 69], and guilt or embarrassment [PR-10, p. 19], and stress. [ST-17] indicates that excessive exposure to dark patterns can irritate and annoy people. As such, we posit that the harm of Emotional Distress appears to stem from a combination of several dark patterns (p. 69). These types of perceived feelings of annoyance, resignation and disengagement are quite consequential, as Conti explains:

> as interface generated annoyance increases, the user will start to become dissatisfied, eventually reaching a point where the user will not accept further frustration in order to accomplish their task. We dub this point of parity the tolerance threshold. Below this threshold the user will remain and above the threshold the user will seek task accomplishment from a competitor, if available [ST-6, p. 4].

The author believes that malicious interface designers seek to operate in a sweet spot just below the tolerance threshold in an attempt to maximise revenue while still retaining the majority of their users. Resignation emerges as either part of the Web experience, or a tradeoff that is still tolerable. [ST-18] reports annoyance and resignation, as participants believed it impossible to avoid online manipulation, and acknowledged that the tradeoff (free service) outweighs negative consequences. [ST-2] adds that participants were

somewhat aware of dark strategies employed by brands in order to mislead users, however they accepted it as part of the internet experience. Regarding disengagement, users want to avoid the dark patterns by stopping using a website or app that deployed them temporarily or permanently [ST-17, p. 69; PR-11].

Dark Patterns Emotional distress is related to several and quite different dark pattern's practices, mostly generally related to Nagging and Forced Continuity practices. For example, annoyance, frustration and stress are related to social comparisons, more particularly, Deceptive Celebrity Endorsements and Social Proof [ST-24, p. 11]. User frustration was mostly felt with the installation of applications without permission (resembling Forced Continuity), unnecessary interruptions, difficult to find content [ST-6, p. 4], and techniques that deliberately forced users to wait and view advertisements [ST-6, p. 4; PR-6, p. 13 and 21]. Confirmshaming was mentioned by one study as triggering guilt or embarrassment regarding certain choices [PR-10, p. 19], and Choice Overload triggers dissatisfaction and uncertainty [PR-5, p. 75]. Interestingly, friction may lead to frustration and time loss [PR-5, p. 92]. The context of cookie banners serves as a placeholder for annoyance and resignation. Kulyk and others study confirms that the presence of cookie banners, and the fact their options are cumbersome to use, furthers resignation, instead of giving control back to the users [ST-14, p. 10]. In [ST-25, p. 10], annoyance is mostly correlated to Sneaking (related to hidden costs), followed by Confirmshaming and Nagging. Finally, addictive practices can contribute to emotional distress (increased pressure, stress, poor sleep) [PR-4].

Related harms Vulnerable individuals may face more severe harms, such as Financial Loss and mental health issues. For example, dark patterns could manipulate someone with a gambling addiction into consenting to targeted ads, leading to more gambling, financial loss, and worsened mental health [PR-10, p. 11].

Remarks Differentiating between various subtypes of harm that lead to Emotional Distress can be challenging, mostly related to supporting evidence. [ST-3] underscores that exposure to dark patterns raises concerns about mental health, especially for vulnerable individuals due to uninformed decisions, with particular emphasis on psychological harm to oneself.

3.3.2.2. Addiction (Low-level). **Definition** Addiction as a harm is not discussed by many sources (only 6 out of the 39 analysed) but those sources analyse it at length. [PR-4] focuses on Addiction but, instead of defining it, offers an overview of how it manifests in the online environment, including: '"excessive or harmful internet use", "smartphone addiction", "technological or internet addiction", "social media addiction"' (p. 4).

Dark Patterns Addiction is mentioned often in relation to Attention Capture dark patterns, such as Infinite Scroll or Auto-Play, as these are 'design mechanisms that use rewarding schemes to keep users entertained and spend more time on a service' [ST-21]. Most sources concur in describing Addiction as the result of the combination and interaction of dark patterns with personalisation (especially for targeted advertising) and recommender systems.

Related Harms Addiction is treated by the sources both as a harm in and of itself, and as conducive to other harms, including Financial loss (the most used example of addiction in policy reports is gambling), cognitive harms, as impacting mental and even physical well-being. It is also often discussed as intertwined, almost blurred, with Loss of time

and Loss of autonomy. This aspect particularly emerges from the empirical studies, which connect Addiction with lack of self-control and self-determination, and waste of time. Addiction is indicated as a factor that turns an average consumer into a vulnerable one, which is especially relevant for consumer protection and the AI Act.

Remarks The fact that addiction is almost unanimously treated as the combination of dark patterns, personalisation, and recommender systems adds complexity from a regulatory perspective, as the phenomenon might fall within the scope of multiple European secondary tools, from data protection to the DSA and AI Act.

3.3.2.3. Labour and cognitive burden (low-level). **Definition** Labour and Cognitive Burden is discussed in 6 out of the 39 analysed sources, albeit briefly. This harm encompasses waste of time, energy and attention [ST-19], the cognitive and perhaps physical effort of users. One study referred to this harm in the context of cookie banners, where users needed to take at least two clicks to opt out of consent, affirming that labour costs impact users' privacy decisions [ST-23]. Cognitive Burden might result from attempts at exploiting consumers' inertia or limited willpower, attention span or time [PR-0, p. 25].

Dark patterns This harm can be found in cookie banners, Nagging, difficult to cancel subscriptions, tricky questions, and Interface Interference practices of online interfaces.

Related harms Cognitive Burdens associated with dark patterns can lead to Loss of Time, and Mental Health Harm, as there is a risk of developing recognised disorders, like internet and gaming addictions, animated by traditional symptoms of dependency, such as excessive use, withdrawal, and tolerance. Cognitive Burdens can also lead to other detriments, such as loss of opportunity, as the attention is bound somewhere else [PR-8, p. 90].

Remarks This harm can be related to both non-material but also to material harms, since it leads to a measurable and tangible Loss of Time. To effectively measure the Loss of Time, it's essential to consider a combination of quantitative and qualitative factors while also acknowledging the subjective nature of time perception.

3.3.2.4. Attention harm (low-level). **Definition** Only 4 studies address this harm explicitly. Attention Harm is understood to be the redirection of the individual's cognitive resources causing them to selectively focus on stimuli from their environment, and it is perceived to be a tradable commodity, whereby users 'pay' for a service with the time they spend on it. Of note, [ST-22] performs the first systematic overview of attentional harms across various domains and proposes the concept of Attention Capture Damaging Patterns (ACDP) defined as 'recurring patterns that designers adopt to manipulate users into spending attention in ways that often lead to a loss of sense of control and time, and feelings of regret.' This harm occurs within social media, video games, and streaming platforms.

Dark patterns Related dark patterns practices are Infinite Scroll, Fake Social notifications, Roach Motel, Disguised Ads and Recommendations, and Autoplay.

Related harms Attention Harm is very closely related to Loss of Time, Loss of Autonomy, Addiction, and indirectly related to Loss of Privacy, depending on the context.

Remarks Considering the resources of time and attention, Attention Harm could be classified as a non-material harm, though with potential material repercussions.

3.3.2.5. Loss of time (meso-level). **Definition** Loss of Time is predominantly discussed in policy reports (specifically 6), while only 2 empirical studies discuss the effects of certain dark patterns on the time spent on a website or app. The harm is not explicitly defined, although it is mentioned by empirical sources in connection with how much time was spent on a task, or engaging with a product/service.

Dark Patterns While some sources only mention Loss of Time as a general undesired effect of dark patterns, others report experimental data showing that certain dark patterns function as an obstacle, forcing users to spend more time before they are able to perform the action they intended to. This is particularly the case of Interface Interference and Sneaking patterns that require users to click more or look for the necessary information over and over again, before being able to move on [PR-4; PR-8]. As a result, users waste time and might eventually give up. In the case of Interface Interference and Sneaking, therefore, Loss of Time has an involuntary negative impact on users. Other dark patterns, however, have 'wasting time' at their very core. Attention capture patterns, such as the Infinite Scroll pattern used by social networks [ST-21; PR-4], make users lose the perception of time passing, and therefore stay longer on the platform. In this permutation, Loss of Time blurs with another harm, Addiction, as they feed each other in a vicious circle and can negatively affect the mental health and psychological well-being of users, especially teenagers [PR-4].

Related Harms Some sources include Loss of Time in the list of moral damages deriving from dark patterns. Others characterise the waste of time as both an undesired effect in and of its own, and as leading to other harms, such as privacy or financial ones, in the case users, by giving up or by experiencing a sunk cost effect, decide to buy a product at a higher price, or share more data than intended.

Remarks In [PR-3], the Dutch ACM recommends expanding consumer protection by recognising non-material harms in a more explicit manner, including Loss of Time. In practice, this idea might clash with the high threshold existing in some Member States for the redress of moral harms (see section 5). Conversely, recognising Loss of Time as a harm seems more likely in the case of Attention Capture Damaging Patterns, because it can be merged with the additive effect and therefore pass a seriousness threshold. It is also possible that Loss of Time leads to a material loss, in which case it can be accommodated more easily also under consumer protection.

3.4. Collective material harms

3.4.1. Weaker or distorted competition (meso-level)

Definition Harms to Competition have been mentioned in just over a quarter of analysed documents (10 out of 39 sources), with 6 of them derived from policy documents. Harms to Competition broadly refer to 'impairing fair competition' [PR-3], 'weakening competition' [PR-11] or 'diminishing competition' [PR-3]. While none of the documents contain a specific definition of harms to competition, the discussion related to competition harms is usually high-level, identifying circumstances in which dark patterns can harm competition. For instance, dark patterns may: generate switching costs to consumers or obstruct their choices [PR-3; PR-8]; weaken or distort competitive pressures [PR-5]; allow certain companies to gain competitive advantages because of deception or making it harder to compare prices and services [PR-11]; extract welfare surpluses

from users [ST-7]; and, ultimately, abuse the position of monopoly power [ST-19]. Therefore, while distortion of competition is a distinct harm, it may also be considered a meso-level harm that encompasses other lower-level categories, such as Reduction of Innovation (see below).

Dark patterns A number of dark patterns may distort consumer behaviour to the extent that leads to weakening or distorting competitive pressures, directly or indirectly [PR-5]. For example, if a platform with significant market power employs the Roach Motel dark pattern, it may lead a high number of users to believe that unsubscribing from a service is not possible, locking-in users and artificially strengthening its market position [ST-9]. Furthermore, Pre-selection, Framing and Obstruction have been used as examples of dark patterns with negative effects on competition [PR-3]. For instance, the antitrust case of *Google Search (Shopping)* has been referenced as an example of anticompetitive self-preferencing behaviour, where ranking and framing of options was used to influence consumer behaviour [PR-5]. Similarly, the UK's Competition and Markets Authority's 'The Mobile Ecosystems Market Study' Interim Report, highlighted how default settings and pre-installations play a critical role in consumers' choice for internet browsers [PR-5].

Related harms The sources discuss Distortion or Weakening of Competition as a specific meso-level harm, which is often discussed in combination with other lower-level harms, such as Reduction of Innovation [PR-8]. Furthermore, competition harms are also related to harms to privacy. The literature illustrates how dark patterns that lead to subversion or obstruction of privacy choice can allow platforms to obtain more consumer data and in turn leverage market effects to strengthen their market position [PR-10]. Finally, dark patterns may result in manipulation that impacts users' decisional privacy and autonomy, which ultimately lead to spending more attention, data and money than intended. This contributes to artificially creating competitive advantages for platforms' deploying dark patterns [ST-7].

Remarks One of the key challenges in identifying Harm to Competition is establishing that a platform with significant market power is engaging in a behaviour that deviates from competition on merits. That entails demonstrating that the platform was able to manipulate users into adopting behaviours that are contrary to their self-interest and in turn reduce their welfare [ST-7]. However, the line between harmful manipulation and legitimate persuasion is thin [ST-7], pointing to the question of when the consumer behaves against their best interest *because* of the behaviour of a platform [ST-19].

3.4.2. Reduction of innovation (low-level)

Definition Harms to Innovation have been mentioned in sources 2 times. While neither of the sources provide a clear definition of Reduction of Innovation Harm, they do explain how dark patterns could negatively affect innovation on the market. [PR-5] explains that '[w]here online choice architecture practices distort consumer decision making, businesses may have less incentive to compete on product attributes that benefit the consumer' (p. 14). As such, harmful interface design features may be cheaper to implement than traditional R&D programmes [PR-5]. In turn, platforms may also engage in so-called 'anti-competitive innovation', where they (re)design products to exclude competition [ST-7]. Here, dark patterns as such may constitute a form of harmful innovation.

Dark patterns None of the sources relate specific dark patterns to the Harms to Innovation. However, as discussed above, Harms to Innovation could occur when different types of dark patterns are deployed at a large-scale and become normalised in markets.

Related harms In both sources, the Harms to Innovation were directly linked to the Harm to Competition. At the same time, it is important to stress that as a wide range of dark patterns, in an aggregate, could lead to negative effects on innovation, it also allows us to infer a possibility for a wide range of (individual) harms.

Remarks One of the main challenges outlined by [ST-7] related to the idea that in the context of 'anti-competitive innovation', it is very 'difficult to draw a line when innovation improves prior technology and when it is meant to solely exclude competition' (p. 31).

3.4.3. Price transparency (meso-level)

Definition Harm to Price Transparency has been referenced in only 3 examined sources. Price Transparency is defined as enabling 'consumers to make informed decisions and thus creat[ing] efficient marketplace' [ST-19, p. 11]. Dark patterns that harm price transparency 'hide the true cost of products from consumers and prevent them from comparison shopping' [PR-8, p. 91].

Dark patterns Firms that deploy dark patterns that harm price transparency, in essence, aim to exploit information asymmetry between the firm and the user. For example, Hidden Costs – also known as Drip Pricing [PR-9] – and Price Comparison Prevention impede informed decision-making [ST-19].

Related harms While sources discuss Harm to Price Transparency as a separate category of collective harm, it is generally conductive to other types of harms, including Harm to Competition, Financial Loss, increased time and effort to avoid additional charges and Emotional Distress, to name but a few. In this regard, consider the example of Drip Pricing. The FTC workshop revealed that consumers felt frustrated that when they begin the shopping process, they do not know how much they will be charged at the end [PR-9]. In addition, the practice hurts competition, as firms that reveal their total price from the onset are placed at a competitive disadvantage with firms that set an artificially low initial price to lure consumers in [PR-9]. Thus, such opaque pricing practices may not only lead to potential financial loss because consumers lose out to cheaper alternatives, but also distortion of competition.

Remarks The true impact of dark patterns' harms to price transparency is very challenging to map out. This is because consumers not only act in accordance with their budgetary constraints, but also may spend time and effort in attempting to circumvent unexpected charges.

3.5. Collective harms that can manifest as either material or non-material

3.5.1. Loss of trust in markets (meso-level)

Definition Loss of Trust in (digital) markets is a relatively commonly referred to collective harm (8 out of 39 sources). Mirroring the individual harm of Loss of Trust, the collective dimension of this harm has not been well-defined across the examined literature. Broadly, it manifests through lowering the credibility of companies, which

may lead to hesitation to interact with the company and ultimately disengagement [ST-17].

Dark patterns Loss of Trust in the markets manifests as a result of a wide range of dark patterns, deployed by one single big firm or several firms [ST-9]. Nevertheless, dark patterns could generally be categorised as being born out of asymmetric information [ST-9], thereby leading consumers to feel misled [PR-5], or those born out of coercion, leading consumers to feel pressured [ST-17; PR-11]. When it comes to the former, dark patterns such as False Hierarchy [PR-5], Partitioned Pricing [PR-5] and Scarcity Cues [PR-6] were considered damaging to consumers' perceptions of fairness and trust. An example of the latter is Forced Action [ST-18].

Related harms Loss of Trust in the market may lead to negative effects on competition, since it may hurt firms that engage in fair and legitimate business practices [PR-8; ST-9; ST-19]. In addition, it may be a result of other incurred harms, including Loss of Privacy, Financial Loss, Emotional Distress, or Time and Effort Burdens.

Remarks Loss of Trust is challenging to measure and studies are not consistent on the subject. For instance, while previous research is positive that in the long-term, consumer happiness and trust is likely reduced by deceptive practices, these longitudinal effects require greater scrutiny, as [ST-18] found that it is the most popular e-commerce websites that most likely contained dark patterns.

3.5.2. Privacy (meso-level)

Definition Collective Harms to Privacy were discussed in 3 examined sources. Notably, collective harms to privacy, to a large extent, represent a collective dimension of individual harms to privacy discussed above. In the sources, this collective dimension is mentioned in two instances: Harm to Privacy that leads to reduction of consumer welfare (welfarist perspective) [PR-10] and Harm to Privacy as a public good (public good perspective) [PR-8]. The former describes a situation where subversion or obstruction of privacy choices lead to competitive (data) advantages to a firm [PR-10]. The latter encompasses widespread use of dark patterns that undermine free choice and normalise lower levels of privacy [PR-10].

Dark patterns Collective Harms to Privacy are derived from the widespread presence of dark patterns that also harm privacy on an individual level. According to [PR-10], practices that harm privacy could be categorised into ones that bring about unwarranted intrusion, such as sharing more personal information than they would otherwise, loss of control and autonomy, and create costs of avoiding or mitigating harm. Such dark patterns include *inter alia* Defaults that expose User Data, Privacy Zuckering, Interface Interference and Obstruction, as well as Confirmshaming [ST-19].

Related harms Overall, sources treat the Loss of Privacy as a harm in and of itself. However, harms to privacy are also closely related to other harms, including Financial Loss and Emotional Distress [PR-10], Attentional Harms [PR-3] and Competition Harms [PR-10].

Remarks As highlighted by [ST-19] and [PR-8], Harms to Privacy can be defined and understood through different lenses: welfarist and public good, as well as human rights, and individual autonomy perspectives. This complicates the measurement of harms to privacy, especially when such harms move from individual to collective dimensions.

3.6. Findings of the analysis on dark patterns' harms

3.6.1. Prominence of harms discussed in the literature

The most prominently discussed harms in our literature review are Emotional Distress (15), Privacy Loss (15), Financial Loss (14), Weaker or Distorted Competition (10), Loss of Autonomy (10), Loss of Trust (9), Loss of Trust in Markets (8), Loss of Time (8), Labour and Cognitive Burden (6). In contrast, the less-commented harms are Addiction (5), Attentional Harm (4), Price Transparency (3), Physical/bodily Harm (3), Privacy (as a collective harm) (3). While Attentional harm is mentioned less, according to the literature review, it is a relational harm connected to several other harms (e.g. Loss of Time, Loss of Autonomy, Psychological Detriment, and indirectly related to Loss of Privacy). The same reasoning applies to the harm of Addiction, as it is treated as a harm per se, but also conducive and blurred to other harms (Financial Loss, Labour and Cognitive Burden, Loss of Time and Loss of Autonomy). Accordingly, individual non-material harms form the majority of dark pattern harms.

3.6.2. Some harms can lead, or be related, to other harms

Our study revealed that some harms are related to other harms, or are conducive to other harms.

Interrelated harms: our analysis confirms that harms are relational and interplay with each other, which shows the complexity of dark patterns' harms. For instance,[34] Loss of Autonomy relates to Addiction; Loss of Privacy and Loss of Autonomy are correlated. Labour and Cognitive Burden relates to Loss of Time. The fact that these harms are related can augment their consequential impact at individual and aggregated levels, requiring comprehensive approaches to address the multifaceted nature of harm, considering also the context wherein they occur.

Conducive to other harms: Sometimes, one harm leads to another harm. Notable examples of *conducive harm* are[35]: individual Loss of Trust, potentially leading to the collective one (loss of trust in the market). Importantly, Loss of Autonomy, which is mostly mentioned in the sources in relation to privacy and e-commerce, seems to lead to a range of other harms. As stated by Mathur et al, '[t]he vast majority of dark patterns attempt to undermine individual autonomy … Concerns about autonomy echo through nearly entire dark patterns literature' [ST-19, p. 13]. In other words, since the mechanisms that underpin dark patterns are manipulative, dark patterns in their nature, infringe user autonomy to varying degrees: this infringement represents a harm in and of itself (Loss of Autonomy) and can lead to other harms, such as Financial Loss.

The interdependence of different types of harms is underscored by the fact that some dark patterns harms were not treated as harms in and of itself, but instead as an *effect* or a *byproduct* of dark patterns leading to other harms. For example, Addiction harm is intertwined, almost blurred, with the Loss of Time and with Loss of Autonomy. Loss of Trust is not treated as a harm in and of itself, but as an effect of dark patterns, which is conducive to other harms, such as Loss of Trust in the Market. From our analysis we infer, therefore, that some harms likely occur *in combination* with other harms.

[34]Additionally, Loss of time relates to Addiction. Attentional harm links to several other harms, such as Loss of time, Loss of autonomy, and indirectly relates to Loss of privacy, depending on the context. Financial loss relates to Addiction.
[35]Additionally, Addiction leads to other harms, such as Financial loss, Cognitive harm; and Loss of time, leads to Privacy or Financial harms.

3.6.3. Harms manifest in different levels of intensity and severity

The analysed sources reveal that dark patterns' harms may manifest in different levels of *intensity and severity*. For example, Privacy Loss can escalate and lead to more extensive processing about user's behaviour, ultimately resulting in targeted advertising or profiling [PR-10, p. 10]. In this line, [PR-6, p. 10] proposes the idea of a 'spectrum of harm', which recognises that some dark patterns manifest as annoyances that are generally viewed as part of users' experiences online, while on the other end of the spectrum, users can suffer direct harm as a result of specific dark patterns. The issue of intensity and severity of harms is also reflected in other sources [ST-17]. Our studies identified specific parameters that influence the intensity of harms: *type* of dark pattern, *context*, *cumulative effects* of dark patterns, and *user vulnerability*, as discussed below.

Dark pattern type In the example of the Loss of Trust, once users sense they have little to no choice left because of a dark pattern (e.g. Forced Continuity) and lose a concrete resource (e.g. privacy, money), they experience Loss of Trust more intensively. Conversely, dark patterns that are less forceful, e.g. Scarcity and Activity Claims, could be considered as a normalised part of users' online experiences and have less impact on reducing trust.

Contextual factors Some contextual factors are considered to play a role in the intensity and severity of (experienced) harms. For example, in regards to Loss of Trust, it could be relevant to discern whether the deception comes from a company or a public authority, the cognitive predisposition of consumer, and the reason for which the dark patter was deployed (for example, to nudge individuals towards a healthier habit, versus interfering with their financial decisions) [PR-5]. Another factor is the *degree of covertness* of dark patterns [PR-11, p. 5]. We infer that covert dark patterns may be so subtle that the user may consider them as a normalised part of experiences online, rendering harms such as Loss of Trust or Emotional Distress when discovered. By the same token, it may also lead to intensified harms to, among others, Attention and Addiction.

Cumulative effects of dark patterns Cumulative effects of dark patterns refers to the situation where a combination of dark patterns may have a compounding effect, 'increasing the impact of each and exacerbating the harm to consumers' [PR-9, p. 5]. For example, Emotional Distress appears to stem from a combination of several dark patterns and, in itself, is composed of several elements.

User vulnerability A combination of (individual) user characteristics may play a role in the intensity of experienced harm. Vulnerability has been referred to in several sources [PR-11; ST-3], and includes characteristics such as age and disability. For instance, [ST-3] demonstrated that 'older generations are not only less able to recognise manipulative attempts, but they are also less aware that their choices and behaviour can be influenced' (p. 733). Vulnerable individuals may face more severe harms, such as Financial Loss and Mental Health Issues [PR-10, p. 11]. Furthermore, younger populations are considered to be more vulnerable to psychopathological developments and harmful behaviours, and mental health conditions established in childhood [PR-11]. Finally, as discussed with regard to the Addiction harm, the underlying behavioural mechanisms of dark patterns, also relevant in context of personalised practices, may turn an average consumer into a vulnerable one.

3.6.4. Dark patterns' modalities may impact harms

Several of the examined sources make a reference to the different modalities of dark patterns, including the use of recommender systems, virtual reality and Internet of Things (IoT). Different types of harms, such as Financial Loss, can occur because of dark patterns embedded across different modalities, but they are less explored in the literature. In this regard, one study underscores the unique implications of dark patterns in IoT [ST-12]. Furthermore, personalisation is mentioned in regard to more advanced forms of dark patterns, with sources making a reference to 'second-generation dark patterns' [PR-8] and 'hypernudging.'[36] According to the literature, Loss of Trust and Addiction harms often go hand in hand with personalised user interfaces, such as recommender systems, leading to regulatory complexity. Personalisation and dark patterns often compromise user autonomy and privacy.

3.6.5. Some harms are both material and non-material

Our taxonomy of harms distinguishes between material and non-material harms, which is in line with the Western legal tradition concerning civil liability (e.g. contractual/tort liability), as referred to in section 2.2. The analysis of each harm and their associated challenges reveals that certain harms can simultaneously fit into both material and non-material high-level categories. For example, Psychological harms are intrinsically non-material in nature, though they can result in or can be associated with material consequences. Annex C depicts our taxonomy of harms, and these hybrid harms are coloured blue. We posit that any endeavours towards their (in)tangibility classification cannot be rigid, as it will fail to capture their multifaceted dynamics, potentially oversimplifying and misrepresenting the harms of dark patterns. Dark patterns's harms classification should be mindful of the complexities that harms entail, for example, the contexts they are identified in, their relational nature to other harms, the types of dark pattern practices that trigger them (either in solo or in combination with other practices), and their different levels of intensity and severity.

3.6.6. Some harms are individual but can lead to collective harms

In our taxonomy, the harms of Privacy Loss, Loss of Autonomy, and Loss of Trust are explicitly categorised as individual harms, though studies indicate that these can lead to collective harms. Moreover, from a broader perspective and to a large extent, collective harms are a manifestation of widespread individual harms, and in this line, it could be inferred that some widespread harms to individuals are conducive and precipitate to other, collective, harms taking place. In this regard, consider individual harms related to Attention, Time Loss and Cognitive Burden. Once these harms manifest on a large scale, they may lead to other collective harms, depending on the context they occur in e.g. dark patterns that impose cognitive burden and lead to friction in their decision-making, may lead to privacy losses and even harm to competition.

[36][PR-8, p. 20], in reference to Marjolein Lanzing, '"Strongly recommended" revisiting decisional privacy to judge hypernudging in self-tracking technologies' (2019) 32(3) *Philosophy & Technology* 549. See also: V Morozovaite, 'Hypernudging in the Changing European Regulatory Landscape for Digital Markets' (2023)15(1) *Policy & Internet* 78.

3.7. Challenges in the classification of harms

3.7.1. Representativeness of harms

Our taxonomy of harms is not intended to be exhaustive. However, we have reached a point of saturation, which means further literature review has not revealed additional harms.[37] Given that dark patterns represent a dynamic field, and there is an increasing attention to the harms deriving therefrom from both regulatory and empirical studies perspectives, it is likely that new harms may be identified and classified in future research. These new harms (or combinations thereof) can then be incorporated into, and expand our current taxonomy. The studies consulted, whether they were policy reports or empirical literature on the harms of dark patterns, were primarily focused on consumer, privacy, and competition issues. While these studies provide a solid representation of the harms of our taxonomy, they can also present a biased perspective, as they concentrate mainly on these specific areas and may neglect other important aspects or broader impacts of dark patterns.

3.7.2. Terminology classification challenges

In our source review we found terminological classification issues across studies, including a lack of consistency in naming certain harms, the use of 'effect', 'consequences', 'negative consequences' or 'impact' to describe harms, and inferred harms, as explained below.

Lack of consistency. When mapping the harms we immediately noticed a lack of consistency in the denomination and terminology used by the various HCI sources and policy reports. This lack of uniformity became even more apparent once we started clustering identical or equivalent harms together. Privacy loss is exemplary in this sense: of the 15 sources discussing it, very few used the same terminology to refer to it. The words used span from 'loss of privacy' to 'unwarranted intrusion', 'obstruction' or 'subversion of privacy choices', 'limited control over data', 'privacy harm', 'invasion of privacy', and so on. In all cases, and as shown by Privacy loss, the presence of recurring keywords (privacy, data) allowed us to form conceptually coherent groups. We ensured that the three authors unanimously agreed before assigning each harm to a group cluster.

Use of various terms to refer to harm. Still from a terminological perspective, it should be pointed out that some sources take a precautionary approach, preferring words like 'effect', 'consequences', 'negative consequences' or 'impact' to describe the harm. This is particularly the case of empirical studies, performed by both scholars and policy makers, which comes as no surprise, since their objective is to identify, measure, or quantify the impact of a certain dark pattern on the study participants's choices or decision-making. The term 'consequences' comes associated with different descriptors, such as negative consequences [PR-6, p. 49; ST-18 p. 190; PR-11 p. 26], harmful consequences [PR-4 p. 5], unanticipated societal consequences [ST-19 p. 11], and also simply consequences, without any specific connotation [ST-9; ST-7 p. 34]. The studies do not clearly indicate whether the varying descriptors for consequences (harmful, negative, or without a specific descriptor) are intentionally used to define a certain degree of severity of the harms caused by dark patterns, or if they result from differing semantics aimed at

[37] BG Glaser and AL Strauss, *Discovery of Grounded Theory: Strategies for Qualitative Research* (Routledge, 2017); B Saunders, J Sim, T Kingstone and others, 'Saturation in Qualitative Research: Exploring its Conceptualization and Operationalization' (2018) 52 *Quality & Quantity* 1893.

describing the adverse impacts of dark patterns. The fact that the same phenomenon (e.g. the loss of time) was referred to by some sources as an 'effect' (wasting time on a website) and explicitly as a 'harm' by others (Loss of time) seems to confirm that, regardless of the apparent neutrality and prudence of the adopted terminology, or difficulty to streamline it, they are still considered negative and undesired outcomes from a normative perspective. For example, this is how one study refers to the self-reported effects of dark patterns: 'All participants described the negative effects of manipulative design on their peers and family. The effects would sometimes reach them, including scams, deceptive designs, addiction, time waste, money loss, insecurities, depression, and social comparison' [ST-24]. Additionally, some sources also use the terms 'loss' and 'damage'. These terms are not synonyms from a civil law perspective,[38] but the EU legislator, international stakeholders, and scholarship often use them either together or interchangeably, which increases confusion and legal uncertainty.

Inferred harms. Finally, some harms were *inferred* from a report or study, whenever these did not expressly mention either harms or damage/losses. For example, in the case of 'Loss of Trust', one source did not expressly mention the keyword 'trust', but the participants of the study had indicated that the presence of certain dark patterns on an e-commerce website gave them the impression of dishonest and manipulative communication. Based on the words used and on the context of the study, we decided that the effect belonged to 'Loss of trust'. These decisions were taken when the authors had reached an agreement.

4. Assessing dark patterns' harms under EU Law

Having mapped the dark patterns' harms emerging from HCI scholarship and policy reports, in this section, we discuss how the harms portrayed in our taxonomy align with EU law requirements. Pursuant to it, we will define and assess harms in the context of data protection (section 4.1), consumer law (section 4.2), competition law (section 4.3), and new laws (section 4.4). We further discuss the challenges of such assessment (section 5). As a research team composed of three scholars, we leveraged our collective experiences in different branches of law. Specifically, our team included established dark patterns scholars, including two with a focus on human–computer interaction (Santos, De Conca), one with experience in consumer law (De Conca), two with a background in data protection law (Santos, De Conca), and one with a focus on competition law (Morozovaite). Across these legal interdisciplinary perspectives, we deployed the legal doctrinal research method to extract the legal requirements relative to harms from legislation and case law. We used a comparative angle to compare requirements across the main relevant branches of EU legislations (consumer, data protection, competition law and new legislation).

Before delving into our legal discussion, it is important to recognise that harms, and especially non-material harms, can be hard to identify (and measure) due to their inherent

[38]C Von Bar, 'The Notion of Damage' in AS Hartkamp, Martijn W Hesselink, Ewoud H Hondius, Chantal Mak and Perron du CE (eds), *Towards a European Civil Code* (Wolters Kluwer Law & Business 2011); Von Bar et al (n 17) 550 and 559. For the purpose of clarification of these concepts, we cite the definitions therewith afforded: 'Damage' means any type of detrimental effect; 'Damages' means a sum of money to which a person may be entitled, or which a person may be awarded by a court, as compensation for some specified type of damage; 'Loss' includes economic and non-economic loss. 'Economic loss' includes loss of income or profit, burdens incurred and a reduction in the value of property. 'Non-economic loss' includes pain and suffering and impairment of the quality of life.

characteristics and the fact that different users may be harmed by the same dark pattern in different degrees. As a result, establishing a *counterfactual scenario* outlining how users would have been impacted by a specific user interface but for the dark pattern is challenging. This is particularly the case with showing *a direct causal link* between online services and the negative consequences these might have on users and societies,[39] raising causality challenges for potential claimants. It thus can be difficult to present *conclusive proof*[40] of harm. As a result, the negative consequences on users can be difficult to assess and quantify; and they are primarily observed by external parties in the aggregate or over the long term.[41] The evaluation of harm, requiring evidence of harm and causation (or at least of their likelihood, as is the case under the UCPD), will remain, to some extent, casuistic due to the fact-intensity of these issues.

Harms to rights might *not always be readily visible or apparent*, for example, when there is invisible data processing or black-box technologies.[42] Their probabilistic nature hampers their evaluation since individuals may incorrectly assess the likelihood of harm due to incomplete information, limited information ('bounded cognitive ability'), and behavioural biases[43] that impair their free, autonomous choices and decisions. For example, Privacy Harms may manifest long after the event leading to it, and hence be undervalued due to 'present bias'.[44] AI harms tend to be systemic and societal, occur at scale, and may constitute risks of future harm rather than current vested harm.[45] Harms are also *diffuse* – the harm to an individual may not be substantive, but when *aggregated,* may lead to significant collective harm. Or, conversely, many operators and events may inflict small insignificant harms on an individual that aggregate to a significant harm for which it is difficult to attribute responsibility or causality. Alongside, harm v*aries and is subjective or contextual* to each person that values one right (e.g. autonomy) to different extents, so the same event and consequences may have different impacts on different individuals, depending on whether or not someone is vulnerable. Harms can be *difficult to avoid* due to economic circumstances such as market power or barriers to switching. In summary, the challenge in identifying harm is due to both intrinsic (subjective, contextual) and extrinsic (economic issues, diffusion, in-visibility, causality) features of dark patterns.

4.1. General data protection regulation (GDPR)

4.1.1. GDPR requirements for compensation of damages

The concept of harm in the GDPR is directly linked to the concept of damage in relation to the processing of personal data. In particular, harm may result from infringement of a GDPR

[39]Access Now 'Towards Meaningful Fundamental Rights Impact Assessment under the DSA' (September 2023) 8 https://www.accessnow.org/wp-content/uploads/2023/09/DSA-FRIA-joint-policy-paper-September-2023.pdf accessed 10 September 2024.

[40]C-460/09 P *Inalca SpA — Industria Alimentari Carni and Cremonini SpA v European Commission* [2013] EU:C:2013:111, para 104; C-150/03 P *Chantal Hectors v European Parliament* [2004] EU:C:2004:555.

[41]Access Now (n 39).

[42]Information Commissioner's Office 'Overview of Data Protection Harms and the ICO's Taxonomy Information Commissioner's Office' (April 2022) 4 https://ico.org.uk/media/about-the-ico/documents/4020144/overview-of-data-protection-harms-and-the-ico-taxonomy-v1-202204.pdf accessed 10 September 2024.

[43]T Mildner, A Inkoom, R Malaka and others, 'Hell is Paved with Good Intentions: The Intricate Relationship Between Cognitive Biases and Dark Patterns' (2024) https://arxiv.org/pdf/2405.07378 accessed 10 September 2024.

[44]Information Commissioner's Office (n 42).

[45]ME Kaminski, 'The Developing Law of AI: A Turn to Risk Regulation' (2023) *University of Colorado Law Legal Studies* Research Paper No.24-5 1, 3; Information Commissioner's Office (n 42).

provision. Article 82(1) provides the right to compensation, establishing the entitlement of 'any person who has suffered material or non-material damage as a result of a GDPR infringement to receive compensation from the controller or processor for the damage suffered'. Member State courts are ultimately responsible for awarding redress to the persons that suffered damages, according to the principle of procedural autonomy (Article 82(6)).

Type of damages The GDPR, in Recital 75, acknowledges physical, material and non-material damages resulting from the processing of personal data**.** Material damages that can be claimed by the data subject may be financial losses, which can occur, e.g. if the data subject becomes the subject of identity theft or fraud due to a GDPR infringement. Non-material damages can comprise personal disadvantages, such as discrimination or damage to reputation (see Recital 85), and depend on the impact on the data subject in the individual case. It is worth noticing that the damage should be *actual* so as to prevent speculative claims.[46]

Requirements for the right to compensation The right to compensation depends on three cumulative conditions: (i) a controller's infringement of the GDPR, (ii) the existence of the damage suffered by the data subject, (iii) a causal link between the damage and the infringement, in line with the standard causal requirement applied in EU Member States. Regarding point (i), the GDPR remedies model introduces a fault-based liability, wherein any organisation involved in unlawful data processing is responsible for that processing, but with an easened burden of proof for the data subject, thanks to a rebuttable presumption of fault of the controller.[47] It suffices that the data subject can prove that a breach of the regulation has occurred on the part of the controller, and that this breach has resulted in eligible damages.[48] A data subject is, therefore, responsible for demonstrating that the breach in question is relevant for, or has caused, the harm suffered, while the controller can demonstrate that the harm is in no way ascribable to its conduct.[49]

Damage threshold assessment. Regarding possible severity thresholds,[50] it is important to consider recent case law in which the CJEU has been asked to determine the difficult borderline between intangible damage and 'mere upset'. While these rulings permit multiple interpretations,[51] they set relevant criteria for the evaluation of GDPR damages. The most relevant for our purposes are:

[46]Case C-300/21 *Österreichische Post* [2023] ECLI:EU:C:2022:756 para 58.
[47]C-667/21 *Krankenversicherung Nordrhein* [2023] ECLI:EU:C:2023:1022; C-687/21 *BL v MediaMarktSaturn Hagen-Iserlohn GmbH* [2024] ECLI:EU:C:2024:72.
[48]J Chamberlain and J Reichel, 'The Relationship Between Damages and Administrative Fines in the EU General Data Protection Regulation' (2019) *Mississippi Law Journal* 667.
[49]Ibid; Case C-300/21 *Österreichische Post* (n 46).
[50]Severity means the magnitude of the risk or its impact if it materialises. Center for Information Policy Leadership, 'Risk, High Risk, Risk Assessments and Data Protection Impact Assessments under GDPR. CIPL GDPR Interpretation and Implementation Project' (21 December 2016) 26 https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_gdpr_project_risk_white_paper_21_december_2016.pdf accessed 10 September 2024. It essentially depends on the prejudicial effect, or the level of consequences of the potential impacts. The gravity/seriousness of prejudice to a human right is usually assessed according to the following three elements: (i) its intensity, (ii) the consequences of the violation, and (iii) its duration, where the intensity of the violation is related to the importance of the violated protected legal interest. See also: S Altwicker-Hamori, T Altwicker and A Peters, 'Measuring Violations of Human Rights: An Empirical Analysis of Awards in Respect of Non-Pecuniary Damage Under the European Convention on Human Rights' (2016) 76 *Zeitschrift für ausländisches öffentliches Recht und Völkerrecht (ZaöRV)/Heidelberg Journal of International Law (HJIL)* 1.
[51]V Janecek and C Santos, 'The Autonomous Concept of "Damage" According to the GDPR and its Unfortunate Implications: Österreichische Post' (2024) 61(2) *Common Law Market Review* 531; S Li, 'Compensation for Non-material Damage under Article 82 GDPR: A Review of Case C-300/21' (2023) 30(3) *Maastricht Journal of European and*

(1) *There is no threshold of seriousness* for the damage (material and non-material), which entails that it is irrelevant whether a damage that actually occurred (e.g. loss of personal data) has reached a certain level of materiality, severity or significance,[52] and as such, it is *not* required the tangible nature of the damage or the objective nature of the infringement;[53]
(2) a mere *fear* of future misuse of personal data can constitute non-material harm, due to an event (e.g. a data breach), even if there is no evidence that any misuse has occurred, but it must be verified that that fear can be regarded as *well founded*,[54] and thus, a purely hypothetical risk of misuse cannot give rise to compensation;[55]
(3) the data subject must *prove* she/he suffered harm, however minimal it is.[56]

4.1.2. GDPR requirements applied to dark patterns's harms

Types of damages acknowledged and compensation requirements Any meso – or low-level individual data-related harm (material and non-material) identified in our taxonomy could, in theory, fall among those that can be remedied under the GDPR, and thus entitle the data subject to redress under Art. 82 GDPR. The GDPR, in fact, accommodates both material and non-material harms, without a severity threshold, as explained above. Regarding the type of harm, as mentioned in section 3.4, individual non-material harms (wherein Emotional Distress was mostly cited) consist of the majority of dark pattern harms identified in the literature review; as a parallel, the FRA social empirical research on EU Member States' remedies in the area of data protection noted that the damage from data protection violations are mostly psychological in nature, such as emotional distress, followed by, although less frequently, financial damages.[57] However, what matters for the application of Art. 82, and for the harm to be acknowledged from a GDPR lens, is that the requirements established therein are fulfilled.

Regarding the *damage requirement*, the damage needs to derive *from* the data processing and, likely, from an infringement of a GDPR provision.[58] Accordingly, the damage is

Comparative Law 335; Gunawan et al. (n 10); We expect both the CJEU and national courts to deliver additional case-law clarifying this topic in the coming year.

[52]C-300/21 *Österreichische Post* (n 46), para. 51; C-456/22 *VX and AT v Gemeinde Ummendorf* [2023] ECLI:EU:C:2023:988 para 18.

[53]C-456/22 Gemeinde Ummendorf (n 53) para 17:

> it cannot be considered that, in addition to the three conditions set out in paragraph 14 of the present judgement, other conditions for establishing liability laid down in Article 82(1) of the GDPR, such as the tangible nature of the damage or the objective nature of the infringement, may be added.

[54]C-340/21 *Natsionalna agentsia za prihodite* [2023] ECLI:EU:C:2023:986, para 85:

> (…) where a person claiming compensation on that basis relies on the fear that his or her personal data will be misused in the future owing to the existence of such an infringement, the national court seised must verify that that fear can be regarded as well founded, in the specific circumstances at issue and with regard to the data subject.

C-687-21, *MediaMarktSaturn* (n 47) para 67.

[55]C-687-21 *MediaMarktSaturn* (n 47) para 68.

[56]C-300/21 *Österreichische Post* (n 46); C-456/22 *Gemeinde Ummendorf* (n 53) para 22.

[57]European Union Agency for Fundamental Rights (FRA), 'Access to data protection remedies in EU Member States' (2014) https://fra.europa.eu/sites/default/files/fra-2014-access-data-protection-remedies_en_0.pdf 28, accessed 10 September 2024.

[58]Whether Art. 82 implies that the damage derives from processing or from an infringement of the GDPR is still debated. See AB Menezes Cordeiro, 'Civil Liability for Processing of Personal Data in the GDPR' (2019) 5 *European Data Protection Law Review* 492.

bound to the user's own data processing context, to which studies referring to Privacy Loss are somehow agnostic, due to the limits of the experiments (and their research questions). Moreover, the harms described in the studies do not refer to *actual* harm suffered by affected people, consisting rather of user's perceptions of 'narratives of harm' contextually experienced by recruited participants when exposed to certain practices. They do not include harms caused by a GDPR infringement, nor harms that were brought before a national court for compensation. Thus, the identified harms in the studies cannot be objectively analysed from a strict GDPR side.

Regarding the *infringement requirement*, the studies assessing dark patterns harms do not take into consideration whether the practices to which the participants of those studies were exposed infringe the GDPR (note that not all denominated dark patterns – as we call them – are or were deemed as illegal by regulatory our judicial decisions). Dark patterns practices, even if privacy-related, were used in the empirical studies, as a factual term, and not a legal term (as today exists in the new laws prohibiting dark patterns). Consequently, the causality requirement would also not be satisfied.

Still concerning the infringement requirement, while Art. 82 proves to be relevant for those dark patterns which violate the GDPR, some issues remain. First, neither the EDPB nor case-law have yet clarified with certainty and consistency which dark patterns infringe the GDPR provisions, and to what extent. So far, only two regulatory decisions[59] explicitly mapped GDPR infringements to dark patterns described in the EDPB guidelines [PR-7]. As per Gunawan et al., several Data Protection Authorities' (DPAs) penalties regard dark pattern infringements that are consent-related, and the authors did not encounter national damage claims regarding consent infringements. This absence suggests that the potential for redress in cases of dark pattern's harms remain underutilised.[60] The argument could be made that dark patterns, by deceiving data subjects, are inherently against the fairness principle enshrined in Art. 5 GDPR, and therefore are always in breach of the GDPR.[61] Depending on the more or less expansive interpretation of Art. 82 and fairness principle in relation to dark patterns, the scope of the protection granted to data subjects against dark patterns harm can vary greatly. Secondly, ultimately the compensation of damages from dark patterns harms under the GDPR will be decided by the national courts of Member States, based on their national liability regimes. The lack of uniformity among national regimes will inevitably lead to different results even in the face of similar or identical situations.[62]

It is not possible to tell, *a priori*, whether damaged data subjects would receive compensation for a specific harm among those identified by the taxonomy, because the evaluation can only be made on a case-by-case basis, considering the concrete circumstances.

Damage threshold assessment The lack of *any* threshold for damages would, in theory, give hope to the redressability of data-related non-material individual harms presented in our taxonomy. This is mostly the case for some Psychological Harms, such as a mere fear of future misuse of personal data, Privacy Loss, and Emotional Distress resulting from a data breach. The fear of future misuse (yet again to be connected to a GDPR

[59]Santos and Rossi (n 9).
[60]Gunawan et al (n 10) 188.
[61]De Conca (n 33).
[62]S De Conca and M Infantino, 'A Difficult Puzzle: the Duty of Care in Member States' Tort Law and the GDPR' (2024) *Datenschutz und Datensich* 48, 566–571.

infringement), if well-founded, could be invoked, in as much as evidentiary proof is portrayed, regardless of how minimal it is. Unfortunately, the Court has not offered a clarification of the term 'well-founded', nor criteria to assess when a moral damage is well-founded. It will be up to the Member States' courts to fill the blanks, which, in relation to dark patterns' harms, seems to be very challenging due the difficulty of even identifying them, as discussed in the introduction to section 4.

4.2. The Unfair Commercial Practices Directive (UCPD)

A prominent objective of the UCPD,[63] in line with the entire EU consumer *acquis*, is protecting consumers against *harms to their economic interests* (Recital 10). The UCPD lends itself to different interpretations. Some experts, applying economic theories, define it as the reduction of consumer's surplus, which can change between average and vulnerable consumers.[64] Others see it as having an adverse effect on the capability of consumers to participate in the market.[65] Setting aside this fundamental question, it is important to note that up until 2019 the UCPD did not contain any specific provision giving consumers a right to redress.

The Omnibus Directive,[66] which updated several pieces of legislation of the consumer *acquis*, introduced a new Article 11a, establishing that

> 1. Consumers harmed by unfair commercial practices, shall have access to proportionate and effective remedies, including compensation for damage suffered by the consumer and, where relevant, a price reduction or the termination of the contract. Member States may determine the conditions for the application and effects of those remedies. Member States may take into account, where appropriate, the gravity and nature of the unfair commercial practice, the damage suffered by the consumer and other relevant circumstances.

This redress provision consolidates decades of doctrine and case-law, which had already affirmed that the remedies available to consumers for damages deriving from UCPD infringements would follow the national (contractual or extra-contractual) rules.[67] While this addition solves the debate concerning whether or not individual consumers were even entitled to remedies under the UCPD, it does not clarify what damages fall within the scope of the provision. No additional clarifications are provided by the European Commission in its Guidelines for the interpretation of the UCPD, where the only damages discussed relate to the practice of consumer lock-in, and are limited to financial loss, loss of

[63]Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council ('Unfair Commercial Practices Directive'), OJ L 149, 11.6.2005, p. 22–39.

[64]P Siciliani, C Riefa and H Gamper, *Consumer Theories of Harm: An Economic Approach to Consumer Law Enforcement and Policy Making* (Bloombsbury, 2021).

[65]A Zardiashvili, *Power and Dignity: The End of Online Behavioural Advertising in the European Union* (PhD thesis, Leiden University, 2024).

[66]Directive (EU) 2019/2161 of the European Parliament and of the Council of 27 November 2019 amending Council Directive 93/13/EEC and Directives 98/6/EC, 2005/29/EC and 2011/83/EU of the European Parliament and of the Council as regards the better enforcement and modernisation of Union consumer protection rules (Omnibus Directive), 7–28.

[67]See, for instance: KA Havu, 'Damages Liability for Non-material Harm in EU Case Law' (2019) 44(4) *European Law Review* 492; E Mišćenić, 'The Effectiveness of Judicial Enforcement of the EU Consumer Protection Law', in Zlatina Meškić, Ivana Kunda, Dušan V Popović and Enis Omerović (eds), *Balkan Yearbook of European and International Law 2019, Balkan Yearbook of European and International Law* (Springer, 2019/2020).

time, and loss of data.[68] As another example of possible consumer harms, a 2017 study identifies several dimensions of consumer detriment, including financial, psychological, time loss, and health.[69]

4.2.1. Consumer law requirements to assess harms

As explained, under the UCPD:

(1) non-material harms are not explicitly addressed;
(2) the aim is to protect consumers against harms to their economic interests; and
(3) redress for individual consumer harms is remitted to national contractual/tort liability regimes.

Actual damage and causal link requirements Similarly as per the GDPR, a breach of the UCPD does not automatically mean that the trader is liable for damages: the *actual* existence of the damage and the *causal link* remain necessary requirements for the damage to be compensated, according to the national regimes.[70]

Threshold of severity Importantly, there is a significant lack of harmonisation among Member States when it comes to tort liability, and moral damages. In some countries, such as Germany, redress for moral damages can be subject to a high threshold of severity, and compensation can only be awarded in cases stipulated by the law. In other Member States, such as France, compensation for moral damages can be awarded in a wide range of cases.[71]

4.2.2. Consumer law requirements applied to dark patterns' harms

Some scholarship has pointed out that, unlike competition law, the European consumer *acquis* has not adopted a theory of harm, capable of clarifying when individuals are harmed by unfair or otherwise unlawful commercial practices, and when harm can be expected, following the infringement of a legal provision.[72] This is not a problem in itself because, concretely, the redress of harm occurs at Member States level, and over time the national regimes have developed detailed doctrines through case-law. However, a lack of theory of harms at the EU level could affect the collective dimension of consumer protection against dark patterns, leaving doubts as to when consumers and traders can expect harm to occur in relation to dark patterns, and what the resulting consequences would be in the consumer *acquis*. According to this view, in EU consumer protection (but not only), the harm is mostly reduced to instances of *breaches of the law*. This approach would not be an obstacle to redress, but it presupposes that all possible harms

[68]European Commission, 'Guidance on the interpretation and application of Directive 2005/29/EC of the European Parliament and of the Council concerning unfair business-to-consumer commercial practices in the internal market' (Notice) OJ C 526.
[69]European Commission, 'Study on measuring consumer detriment in the European Union' (February 2017) Final Reports Part 1 – Main report https://commission.europa.eu/document/download/a48e1d90-6728-4e46-bf64-6f0e9f15a935_en?filename=consumer-detriment-study-final-report_en.pdf accessed 10 September 2024.
[70]J Stuyck, 'The Court of Justice and the Unfair Commercial Practices Directive' (2015) 52(3) *Common Market Law Review* 721.
[71]For an overview of the differences among Member States (and beyond): M Bussani, AJ Sebok and M Infantino, *Common Law and Civil Law Perspectives on Tort Law* (Oxford University Press, 2022).
[72]F Esposito, 'Towards a General Theory of Harm for Consumer Law' (2021) 44(2) *Journal of Consumer Policy* 329; Esposito and Sibony (n 19).

are accounted for in the law, that the law is perfect and accurately reflects the reality of the markets and commercial practices.[73] This is particularly significant in the context of harms that manifest online, due to the fast technological developments.

As our taxonomy shows, HCI experts and policy makers have identified a wide range of individual and collective harms, both material and non-material. We do not think that establishing a theory of harm is the only solution to this issue, especially due to the co-existence in consumer law of both individual and collective harms. However, the analysis of the empirical studies shows that at the moment it would be difficult to anticipate if and how some of the harms deriving from a dark pattern, in a certain context, can be compensated. This can be remedied by intervening at the EU level, either by attempting at creating a theory of harms for consumer protection, or by issuing clear guidelines on the nature of consumer harms. Commission guidelines, in this sense, would have the advantage of being faster to draft than, for instance, the development and consolidation of case law by the CJEU, and could include the perspectives of national consumer protection authorities and other stakeholders.

To complicate matters further, the lack of harmonisation among national regimes can result in the same harm being granted compensation in some Member States, but not in others. The biggest difference in this sense regards non-material harms, which are not explicitly mentioned by the UCPD. Therefore, it is not guaranteed that non-material harm will obtain redress under the consumer protection regime. This is why the Dutch Authority for Consumers and Markets suggested explicitly to include, among the possible harms to consumers, some non-material ones, such as 'time lost, emotional harm, privacy lost or addiction' [PR-3, p. 4]. Other sources also suggested adding addictive design patterns among the practices prohibited by the UCPD Annex I, to strengthen consumer protection [PR-8].

Besides individual redress, consumer protection in the EU also serves the purpose of protecting the general interest to fairness in business, which indirectly can prevent market distortions and the loss of trust in the market by consumers. The UCPD tasks Member States with laying down penalties for the infringement of the Directive itself. The Omnibus Directive amended this provision, adding that, in establishing the penalties, Member States must consider, among others, 'any action taken by the trader to mitigate or remedy the damage suffered by consumers' (Article 13(2)(b) consolidated UCPD). While the harm remains unspecified and undetailed, remedial actions against it are supposed to be used by Member States in calculating effective, proportionate and dissuasive penalties.

4.3. Competition law

4.3.1. EU competition law requirements for harms

In free market economies, competition law is central to protecting the market from firms engaging in anticompetitive behaviour: it is a legal field most closely associated with curbing the negative manifestations of firms' market power, which lead to distortion of competition and consumer harm.[74]

[73]Esposito and Sibony (n 19) 20.

[74]In the EU, competition laws are comprised by cartel prohibition, abuse of dominance prohibition, merger control rules and competition rules applicable to public undertakings and those given special or exclusive rights by Member States. See: European Commission, 'Competition Law Treaty Provisions for Antitrust and Cartels: The Treaty on the Functioning

Exclusionary and exploitative abuses For the purposes of this article, the legal analysis of dark patterns' harms is assessed in relation to *abuse of dominance prohibition* enshrined in Article 102 TFEU. Article 102 TFEU concerns unilateral behaviour of undertakings holding a *dominant* position in a specific relevant market. Market power is not considered problematic as such. Instead, the dominant undertaking holds a special responsibility not to abuse its power.[75] In this regard, Article 102 TFEU contains a non-exhaustive list of abusive practices.[76] A distinction can also be made between different types of abuses: exploitative and exclusionary. *Exploitative abuse* relates to dominant firms' behaviour that exploits customers and consumers directly, for instance, through excessive pricing. *Exclusionary abuse* is about dominant firms artificially raising barriers to entry and expansion and in turn excluding competitors from the market. Notably, the recent developments in regard to digital competition has brought forward a new generation of hybrid cases that possess characteristics of both.[77] For a practice to be considered abusive and thus trigger the application of Article 102 TFEU, it is necessary to establish a logically consistent *theory of harm*, which is able to articulate why a specific behaviour causes harm to competition and ultimately consumers.[78] The notion of 'harm' in competition law is not neatly defined, but it generally pertains to harms that occur in the economic domain and relates to different *parameters of competition: price, choice, quality and innovation*.[79]

Collective harms Harms to competition are *collective*, meaning that they are capable of affecting the competitive constraints in a specific relevant market. Ultimately, competition law safeguards consumers by preventing, *inter alia*, reduction of consumer welfare,[80] distortion of the competitive process[81] and competitive structure.[82] It is notable that EU competition law does not seek to protect competitors; it recognises that even powerful undertakings can sustain and expand their market positions through competition on merits and, in turn, push weaker competitors out of the market.[83]

Anticompetitive effects Since the late 1990s, there was a shift from the form-based to effects-based approach to Article 102 TFEU enforcement, meaning that a practice is not considered per se unlawful and instead it is necessary to build a theory of harm and

of the European Union (TFEU) Articles relevant to Competition law' https://competition-policy.ec.europa.eu/antitrust-and-cartels/legislation/competition-law-treaty-articles_en accessed 9 June 2024.

[75]C-322/81 *Nederlandsche Banden Industrie Michelin (Michelin I) v Commission* [1983] ECR 3461, para 57.

[76]C-6-72 *Europemballage Corporation and Continental Can Company Inc. v Commission of the European Communities* [1973] ECLI:EU:C:1973:22, para 26.

[77]P Ibanez Colomo, *The New EU Competition Law* (Hart Publishing, 2023) 134.

[78]H Zenger and M Walker, 'Theories of Harm in European Competition Law: A Progress Report' in Jacques Bourgeois and Denis Waelbroeck (eds) *Ten Years of Effects-based Approach in EU Competition Law* (Bruylant 2012).

[79]European Commission, 'Guidance on the Commission's Enforcement Priorities in Applying Article 82 of the EC Treaty to Abusive Exclusionary Conduct by Dominant Undertakings' (Communication) OJ C 45/7, para 5. Note, 'generally' because some harms could also touch upon non-economic harms e.g., C-252/21 *Meta Platforms Inc. v Bundeskartellamt* [2023] ECLI:EU:C:2023:537 drew an implicit connection between data protection harms and harms to competition. See also on the broader discussion on the non-economic goals of EU competition law: A Gerbrandy, 'Rethinking Competition Law within the European Economic Constitution' (2019) 57(1) *Journal of Common Market Studies* 127.

[80]Consumer welfare in the EU refers to consumer surplus, and is understood through the economic efficiencies lens. It relates to parameters of competition: price, choice, output, quality, innovation. See: V Daskalova, 'Consumer Welfare in EU Competition Law: What Is It (Not) About?' (2015) *Competition Law Review* 131.

[81]European Commission, OJ C 45/7 (n 80) para 6.

[82]C-6/12 *Europemballage Corporation and Continental Company Inc. v. Commission of the European Communities (Continental Can)* [1973] ECLI:EU:C:1973:22, para 12.

[83]European Commission, OJ C 45/7 (n 80) para 6.

show that it may lead to *anticompetitive effects*.[84] Nevertheless, the standard for establishing anticompetitive effects is disputed, since the CJEU has consistently held that anticompetitive effects can be actual or potential.[85] While there is no requirement to demonstrate actual effects on competition,[86] the potential effects cannot be hypothetical.[87] Depending on the practice in question, different thresholds may be applied to show that behaviour is anticompetitive.[88]

All in all, establishing Article 102 TFEU infringement necessitates showcasing how a specific practice harms competition and consumers. The establishment of a credible theory of harm requires demonstrating that firms' behaviour deviates from competition on merits and may lead to actual or potential anticompetitive effects.

4.3.2. Application of competition law requirements to dark patterns harms

Exclusionary and exploitative abuses Regarding exclusionary abuses, which constitute the Commission's enforcement priority to date,[89] it is notable that dark patterns are *not* considered an independent type of abuse. In fact, there is no case that explicitly covers dark patterns' harms. However, dark patterns could become *the means* for anticompetitive exclusionary behaviour to take place.[90] For instance, scholarly literature has examined Roach Motel dark patterns, which would lead to higher switching costs for consumers and even obstruct their choices, thereby artificially creating advantages to the dominant firm.[91] Examples include, among others, default (Preselection) and framing interventions that impact consumers' decision-making and could allow dominant firms to engage in problematic self-preferencing behaviour,[92] and Forced Action dark patterns that skew or limit their choices. Importantly, when different dark patterns are deployed in a large-scale, systemic manner, the compounded effects could lead to weakening or distorting competition and reducing the incentives to innovate to the benefit of consumers. Furthermore, manipulating consumers in a large-scale manner may constitute exploitative abuse, as the goal of the practice is to extract as much consumer surplus as possible.[93]

Collective harms When it comes to dark patterns' harms, this article makes a distinction between individual and collective harms. As discussed above, in relation to the public enforcement of competition law, individual harms fall outside the scope of competition

[84]Zenger and Walker (n 79).
[85]C-52/09, *TeliaSonera*, EU:C:2011:83, para 64. See on in-depth discussion about the meaning of anticompetitive effects in EU competition law: Pablo Ibáñez Colomo, 'Anticompetitive Effects in EU Competition Law' (2020) 17(2) *Journal of Competition Law and Economics* 309.
[86]*Michelin v Commission (Michelin II)* (T-203/01) [2003] ECLI:EU:T:2003:250, para 239.
[87]*Post Danmark A/S v Konkurrencerådet* (Case C-209/10) [2012] ECLI:EU:C:2012:172, para 65.
[88]In regards to showing the probability of anticompetitive effects taking place, Ibáñez Colomo makes a distinction between plausibility, likelihood and certainty thresholds. See: P Ibáñez Colomo, 'Anticompetitive Effects in EU Competition Law' (2020) 17(2) *Journal of Competition Law and Economics* 309, 318.
[89]European Commission, OJ C 45/7 (n 80).
[90]V Morozovaite, 'The Future of Anticompetitive Self-preferencing: Analysis of Hypernudging by Voice Assistants under Article 102 TFEU' (2023) 19(3) *European Competition Journal* 410.
[91]See [PR-8, p. 91]; [PR-10, p. 10].
[92]For instance, Google Shopping, to a large extent, was a case about *framing of options* and intentional demotion of competitors' products and services. Google Android was about tying, i.e., setting Google Search app as a default in Google Android's OS, and in turn mobile ecosystem. See: Case T-612/17 *Google and Alphabet v Commission (Google Shopping)* [2021] ECLI:EU:T:2021:763; Case T-604/18 *Google and Alphabet v Commission (Google Android)* [2022] ECLI:EU:T:2022:541.
[93]I Graef, 'The EU Regulatory Patchwork for Dark Patterns: An Illustration of an Inframarginal Revolution in European Law?' in Ramsi A Woodcock (ed), *Toward an Inframarginal Revolution: Markets as Wealth Distributors* (Cambridge University Press 2023).

law application since it is exclusively concerned with collective harms that materialise in the market domain. In this regard, competition law is distinguished from consumer law because it aims to maintain a competitive market environment by sanctioning the behaviour of (powerful) firms, instead of focusing on protecting individual consumers by ensuring fair treatment in their dealings with businesses. In addition, non-material harms generally do not fall within the remit of competition law.

Anticompetitive effects Building a credible theory of harm and establishing actual or potential anticompetitive effects in relation to a dominant undertaking's unilateral behaviour is a resource-intensive endeavour. Drawing a line between a legitimate business strategy and a practice that is capable of producing *anticompetitive effects* is challenging, especially in cases that involve elusive concepts such as manipulation and deception. Furthermore, as discussed above, the use of dark patterns is unlikely to constitute an abuse of dominance on its own, but instead it may be considered to contribute to anticompetitive outcomes. In this regard, authorities show increasing proactivity in including behavioural analyses in their investigations and building credible theories of harm in order to show that behavioural interventions, which could include dark patterns, impact consumers' behaviour and contribute to anticompetitive behaviours that weaken or distort competition.

Enforcement While the above discussion shows that, in principle, collective dark patterns' harms could be addressed by EU competition law, there are practical challenges that may hinder public enforcement in this regard. In particular, while dark patterns can be a widespread issue on the market, competition law requires case-by-case assessment for a very specific set of circumstances – a dominant undertaking engaging in behaviour that distorts competition and harms consumers, not capturing behaviour by smaller platforms, thereby addressing the problem only partially. Since dark patterns in the market domain are, to a large extent, covered by the UCPD and the emergent EU digital regulations (see section 4.4), the role for competition law remains, perhaps pragmatically, a limited one.[94]

4.4. Dark patterns's harms in the new EU laws

For the past forty years, competition law, consumer law, and data protection have been the main legal instruments through which the European Union has developed a regulatory framework for digital technologies and services. In recent years, those three fundamental regimes have been updated and complemented by new secondary legislation which, by virtue of their novelty, contain provisions that explicitly refer to dark patterns and manipulative design: the Digital Markets Act, the Digital Services Act, the Artificial Intelligence Act, the Data Act. These new laws are valuable for our analysis, but difficult to catalogue into an existing branch of European law, as they contain provisions belonging to consumer protection, contract law, competition, data protection, etc. For this reason, we analyse their coverage of harms caused by dark patterns separately.

[94]Enforcement of consumer protection laws may have synergies with competition law. See: Federal Trade Commision, 'FTC Takes Action Against Amazon for Enrolling Consumers in Amazon Prime Without Consent and Sabotaging Their Attempts to Cancel' (June 2023) https://www.ftc.gov/news-events/news/press-releases/2023/06/ftc-takes-action-against-amazon-enrolling-consumers-amazon-prime-without-consent-sabotaging-their accessed 9 June 2024.

4.4.1. Digital Markets Act

The DMA is a pro-competition regulation, providing a set of harmonised rules applicable to companies designated as 'gatekeepers',[95] which offer core platform services such as online intermediation, social media, operating systems, cloud services or search engines.[96] In contrast to European competition law, which necessitates establishing a credible theory of harm to show that a dominant firm's behaviour is anticompetitive, the prohibitions outlined in Articles 5–7 DMA impose *per se* rules on gatekeeping platforms.

While the DMA does not refer to dark patterns or their harms explicitly, it contains several provisions that capture some of the challenges associated with dark patterns. Article 13 prescribes that gatekeepers should not use dark patterns to manipulate or deceive users as a way to circumvent their obligations under Articles 5–7.[97] For instance, Recital 37 stipulates that gatekeepers should not design, organise or operate their online interfaces in a way that deceives, manipulates or otherwise materially distorts or impairs the ability of end users to freely give consent. It concretises that not giving consent should not be more difficult than giving consent (a common dark pattern known as 'Obstruction'), and calls for a user-friendly solution to provide, modify or withdraw consent. Article 5(2)(d) and 5(8) impose a direct prohibition on 'Forced Subscription'. Article 5(2) limits the scope for 'Nagging' in the context of requesting user consent to process, combine and cross-use user's personal data more than once within one calendar year. Article 6(3) also supports the easy change to default settings 'on the operating system, virtual assistant and web browser of the gatekeeper that direct or steer end users to products or services provided by the gatekeeper.'[98]

Regarding in particular, harm, Recital 70, directly related to the anticircumvention provision, does not explicitly refer to harm in the legal provision scoping dark patterns. Instead, it includes a wider formulation that goes beyond the technical circumvention of the legal obligations in the DMA and mentions the protective purpose to ensure autonomous decisions and choices. We posit that the Harm to Autonomy can be indirectly inferred when dark patterns are deployed by gatekeepers. Since DMA is about large-scale practices that impact fairness and contestability of markets, it refers to collective harms that result from these practices. It therefore includes as the subjects of harm end-users and business users.

4.4.2. Digital services act

The DSA sets out a harmonised framework of rules for a safe, predictable and trusted online environment.

Concerning harms, Article 54 entitles the recipients of a service (i.e. any natural or legal person using an intermediary service) to seek redress for *any* damage or loss *deriving from the infringement of the obligations* set out in the DSA itself. Under the DSA, therefore, the scope of redress is limited to those harms (damage or loss) deriving from an infringement

[95] Articles 3(2) and 3(8) DMA set out the quantitative and qualitative criteria respectively, which must be met to qualify as gatekeeper.
[96] Article 2(2) DMA.
[97] For instance, Article 5(2) does not allow combining data sets across different platform services unless user consent is obtained for this purpose. If a gatekeeper used a dark pattern in a consent mechanism to trick users to agree to a combination of data sets, they would be circumventing the obligation within the meaning of Article 13.
[98] For further discussion, see [ST-9].

of the online platforms providers' obligations. This is particularly interesting, since Article 25 DSA expressly prohibits providers of online platforms to

> design, organise or operate their online interfaces in a way that deceives or manipulates the recipients of their service or in a way that otherwise materially distorts or impairs the ability of the recipients of their service to make free and informed decisions.

Consequently, any dark pattern recognised to be in breach of the DSA (specifically of Article 25, but also of any other relevant obligation) can trigger Article 54, if the infringement has caused harm.

The reference to any harm should include both material and non-material harms. However, from the letter of Article 54 and Recital 121, it is clear that this right – and the corresponding liability of the online platforms – is subject to the national liability regimes, with all the consequences deriving therefrom with regard to the lack of harmonisation. This is especially true for non-material harms, as explained above. Like the DMA, the DSA mentions the protective purpose of Article 25, aiming to ensure autonomous decisions and choices. The related Recital 67 confirms that manipulative practices can be used to persuade the recipients of the service to engage in unwanted behaviours or into undesired decisions which have 'negative consequences' for them. Negative consequences, is safe to assume, can be a reference to a variety of possible harms deriving from dark patterns.

It is worth noting that the applicability and scope of Article 25 are currently debated. The second paragraph of the article limits the prohibition to use dark patterns, establishing that it 'shall not apply to practices covered by' the GDPR and the UCPD. The scholarship has argued that the terms 'practices covered by' does not equate 'practices already prohibited under':[99] this could create a loophole for online platforms, significantly limiting the application of the DSA. This circumstance is not clarified by Recital 121, which affirms that the right to seek redress under the DSA is without any prejudice to the right to seek other possibilities for redress under the consumer protection regime. It is not clear whether the two rights could be exercised cumulatively or the rights deriving from consumer protection should prevail. Finally, the scope of Article 54 is further limited by Recital 121, specifying that the liability for damages of the online platforms is limited by the exemption for mere conduit and caching intermediary services provided by Articles 4 and 5 DSA.

Additionally, the DSA goes beyond the individual relationship between an online platform and the recipients of the service,[100] scoping individual and collective rights and demanding the assessment of systemic risks of any 'actual or foreseeable negative effects for the exercise of fundamental rights', as per Article (34)(1)(b). This is also reflected in the fact that recipients of the service (or a body/entity representing them under Article 54 DSA) are entitled to lodge a complaint with the relevant national authorities, to bring their attention to alleged infringements of the DSA (Article 53). These complaints can lead to investigations and possible fines, and can be used by competent authorities to scope systemic risks or cross-cutting impacts.

[99]C Santos, N Bielova, S Ahuja, C Utz, C Gray and G Mertens, 'Which Online Platforms and Dark Patterns Should Be Regulated under Article 25 of the DSA?' (2024) *Dark Patterns & Manipulative Design: Conceptualising and Systematising a Key Contemporary Phenomenon from a Legal Perspective and Beyond* (Edward Elgar, forthcoming).
[100]Inge Graef (n 94) 10.

4.4.3. Artificial Intelligence Act

The AI Act provides for harmonised rules for the placing on the market, putting into service and use of AI systems in the EU. Following a risk-based approach, the AI Act classifies certain uses of AI systems as prohibited, and sets out requirements for high risk and other AI systems. It includes two bans of specific manipulative practices. Firstly, it prohibits AI systems that

> deploy subliminal techniques or purposefully manipulative or deceptive techniques, with the objective, or the effect of, materially distorting the behaviour of a person or a group of persons by appreciably impairing their ability to make an informed decision, thereby causing a person to take a decision that that person would not have otherwise taken in a manner that causes or is likely to cause that person, another person or group of persons significant harm. (Art. 5(1)(a))

Secondly, it prohibits practices that exploit the

> vulnerabilities of a person or a specific group of persons due to their age, disability or a specific social or economic situation, to or the effect of materially distorting the behaviour of that person or a person pertaining to that group in a manner that causes or is reasonably likely to cause that person or another person significant harm. (Art. 5(1)(b))

Both prohibitions refer to individual and collective harms, and present a high threshold (the harm must be significant). The AI Act does not provide any specific remedy to individuals damaged by a violation of the prohibitions contained in Article 5. Recital 170 explicitly states that 'Union and national law already provide effective remedies to natural and legal persons whose rights and freedoms are adversely affected by the use of AI systems.' This solution seems very reasonable, since Article 5 AI Act (like other AI Act provisions) do not grant individual rights, but consist of prescriptions and obligations to regulate the entry on the internal market of AI systems. The only remedy offered to legal or natural persons is the possibility to lodge a complaint with the relevant market surveillance authority, to bring to their attention any infringements of the AI Act. The right to lodge a complaint has a more collective function, as it triggers the market-surveillance powers of the authorities established under Regulation 2019/1020 on market surveillance and compliance of products.[101] Individual remedies can be sought through other existing EU or national laws, including national liability regimes, the GDPR, the consumer *acquis*, or the DSA, where applicable.

4.4.4. Data Act

The DA aims to maximise the benefits of data in the internal market, by enabling a broader range of stakeholders to gain control over the data generated using connected products and related services (such as Internet of Things devices and their apps). The DA contains provisions that enable users (i.e. natural or legal persons who own a connected product or have the right to use them/the related services) to share such data with third parties. The DA contains specific provisions that prohibit data holders and data recipients from making the exercise of the user rights unduly difficult (Articles 4(4), 6(2)(a), Recital 38). It explicitly refers to design that presents choices to users in a non-neutral manner,

[101] Regulation (EU) 2019/1020 of the European Parliament and of the Council of 20 June 2019 on market surveillance and compliance of products and amending Directive 2004/42/EC and Regulations (EC) No 765/2008 and (EU) No 305/2011, PE/45/2019/REV/1, OJ L 169, 25.6.2019.

and to 'subverting or impairing the autonomy, decision-making, or choices of the user via the structure, design, function or manner of operation of a user digital interface'.

The DA does not contain any provision establishing the right to redress for tangible or intangible harms, but establishes the right of users to lodge a complaint with a national competent authority against the infringement of one of their rights established by the DA (without prejudice for the right to recur to administrative or judicial procedures, Arts. 4(3)(a)-(9)(a), 5(12)(a), 17(5)). The right can be exercised individually or collectively. The DA conceptualises the infringement of a right to which the user is entitled according to the DA itself, as a detriment in and of itself. While the harm appears to be individual, the right to lodge a complaint can be exercised collectively. With regard to other possible harms deriving from the infringement of the DA, users retain the right to seek judicial redress, following the national rules for contractual or tort liability.

5. Discussion

At the beginning of this article, we asked the question: *to what extent are dark patterns' harms legally recognisable and enforceable in EU law?* The answer emerging from our research is, in the best (or worst) legal fashion: to a certain extent. Potentially, all the individual and collective harms identified in our taxonomy at meso– and low-level can be recognised and remedied in one or more of the EU laws analysed. However, they might also not. While policymakers and scholarly literature confirm that dark patterns can be harmful both at individual and collective levels, the effective redress of the harms caused by dark patterns faces several challenges, as confirmed by the different legal requirements and thresholds (see Section 4). In this discussion section, we outline the findings stemming from both the literature review and EU law analysis (Sections 5.1–5.5) and provide recommendations for regulators and social scientists (Sections 5.6).

5.1. The inconsistent description of harm complicates the interpretation of harm and redress

We discussed in section 3 that studies face inconsistency issues regarding harm (they make use of inferred harms, 'effect', 'consequences', 'negative consequences' or 'impact' to define harms, and they lack uniformity in harm terminology). Similar challenges arise in the analysis of EU laws addressing the harms of dark patterns. The terminology and language used in multi-pronged regulations is lacking a consistent and aligned use of the concept of harm across various EU laws (see Annex C). EU laws refer to harm implicitly (e.g. AI Act), but the terminology varies, ranging from damage (GDPR), impairing autonomy and decision-making/choice (such as in the DMA, DSA, and AIA), creating undue difficulties in exercising rights (DMA, DA), having negative consequences for consumers (UCPD, DSA), to manipulating and using subliminal techniques (AI Act).

Moving beyond terminological inconsistencies,[102] EU laws offer a patchwork of legal protections against harms caused by dark patterns in terms of *scope*, *users*, and *types of*

[102]Busch and Fletcher (n 32) 32.

harms. Annex B offers a summary of the comparative analysis among EU laws. In fact, EU laws afford different scopes of protection, either individual, collective or both. These cover different types of users, such as end-users and business users; for example, in the context of the DSA and DA, the party entitled to protection against dark patterns can be both business users and consumers (Recital 2). In the GDPR, it is only natural persons (data subjects). In the DMA, the harm is related to the impact that impaired user decision-making can have on the contestability and fairness of digital markets (not users).[103] Finally, EU laws scope different types of harms. For instance, the DMA and DA protect both autonomy loss and the detrimental effect resulting from the impossibility to exercise rights,[104] while the AIA covers significant harm and autonomy loss, and the GDPR and DSA cover any damage (physical, material or non-material damage).

This lack of uniformity even in the letter of the law is, on one side, inevitable, since each secondary legislation has different material scopes and objectives, though the legislator drafting these laws is the very same; on the other side, for a phenomenon as transversal as dark patterns which has the potential to violate all of those laws, it creates more interpretational issues across laws, uncertainty, and is an upstream obstacle to elaborating a clear overview of dark patterns' harms, their regulation and redress under EU law.

5.2. Redress for material harms

Looking at individual material harms deriving from dark patterns, the combination of EU secondary laws and national liability regimes offers private enforcement redress to damaged persons.

Economic loss is probably the simplest to address. Both the GDPR and UCPD provisions are built around the archetype of financial loss, a damage that is monetary, quantifiable, and easily identifiable in its boundaries. Other material harms, such as Physical/Bodily Harm and (potentially) Loss of Time, can give rise to claims under the GDPR or the UCPD, depending on an existing infringement and the factual circumstances and, once again, on the Member States' national regimes. When Privacy Loss manifest as material or non-material damage, it all under the GDPR's regime,[105] though doubts remain on how this harm would be scoped under the UCPD when it manifests as non-material, since the Directive was clearly focused on the economic interests of consumers, and the debate about the inclusion of moral harms remains open.

Additionally, the DSA and DA offer natural and legal persons the possibility to lodge a complaint with the relevant national authorities, and the possibility to recur to the national courts for damage or losses caused by an infringement of the law itself. This means that individuals who suffered material harm from dark patterns will be able, in theory, to seek redress before their national courts, under the applicable national liability regimes.

[103]Idem.

[104]The impossibility to exercise rights (or restrictions thereto) consist of interferences with these rights, even if they can be reconducted to just legal violations. Such restrictions can provoke detrimental effects to end-users, depending on the practice that caused the restriction or the context.

[105]Notably, Privacy loss and Loss of Time could manifest as non-material harms, and in this form they could still be recognised under the GDPR.

5.3. Redress for non-material harms

Redress for non-material harms is less certain and more fragmentarily captured, as noted time and time again. The only law explicitly mentioning them (and supported by a body of CJEU case-law) is the GDPR, with the conceptual and procedural challenges presented in section 4.1.2. Within the scope of the GDPR, the proposed harm-based taxonomy suggests that, in the context of data processing, especially the harms of Privacy loss, Loss of Trust, and Psychological detriment may be eligible for redress.

There is a higher likelihood of securing compensation in cases involving the harm of Loss of Autonomy. The recently enacted EU laws, including the DSA, DMA, DA and AIA, emphasise the purpose of the protection of autonomous, free, and informed choices and decisions in their provisions targeting dark patterns,[106] 'moving the concept of personal autonomy from the position of a meta-principle to the position of an explicitly protected value'.[107] These laws also signify to the Loss of Autonomy in several ways (see Annex B), e.g. 'impair user autonomy'; 'unwanted, undesired behaviours or decisions'; 'not in the recipient's interests'; 'causing the person to take a decision that that person would not have otherwise taken'. The legislative focus on autonomy as a policy objective aligns with the observations made in the foundational study of Mathur et al. [ST-19] and others[108] that posit that autonomy underpin most of the mechanisms of dark patterns, and with the qualification of Loss of Autonomy as a type of harm in the literature review (see section 3.1). The EU's legislative efforts appear to be well-founded and essential for preserving user autonomy from harmful dark patterns. The redressability of autonomy was not yet discussed in either EU-based guidelines nor case law, but one can reason that such harm could be recovered once it is instantiated in *related* concrete consequential harms (e.g. Loss of Control, Anxiety, or other mental, physical, economical harms) which might entail that it might not be a harm per se, but a harm that *leads* to other harms.

Regarding other laws, it is not certain that a damaged individual can bring a claim for non-material harms (with the exception of Loss of Autonomy) under the UCPD, DA, and AI Act. Even if that question is answered positively, ultimately the possibility for redress is dictated by national courts and national tort liability regimes. What is striking, however, is that individual non-material harms constitute the majority of dark pattern harms. This means that, with the current legislative landscape, the losses that are more prominent in empirical studies and more discussed among experts, are also those that the law cannot accommodate with certainty and with clear requirements.

[106]The link between consumers' ability to make informed decisions and the principle of consumer autonomy is made explicit in these EU legislative acts. As Brenncke poses, specifying the autonomy violation for the purposes of these laws relies on differentiating between autonomy-violating and non-violating external influences on consumer decision-making processes. M Brenncke, 'A Theory of Exploitation for Consumer Law: Online Choice Architectures, Dark Patterns, and Autonomy Violations' (2024) 47(1) *Journal of Consumer Policy* 156.

[107]M Gartner, 'Regulatory Acknowledgement of Individual Autonomy in European Digital Legislation: From Meta-principle to Explicit Protection in the Data Act' (2022) 8(4) *European Data Protection Law* 462 https://edpl.lexxion.eu/article/EDPL/2022/4/6 accessed 19 September 2024.

[108]S Ahuja and J Kumar, 'Conceptualizations of User Autonomy within the Normative Evaluation of Dark Patterns' (2022) 24(52) *Ethics and Information Technology* 51; Brenncke (n 107).

5.4. Redress for collective harms

The responsibility largely falls under EU competition law, which could, in principle, address some collective material harms of dark patterns, including Weakening and distortion of competition as well as Reduction of Innovation. While individual material or non-material harms do not fall within the scope of competition law, it is important to underscore that several individual harms, once manifested in a large-scale manner, could contribute to collective material harms. For example, Loss of time and Attention, as well as collective Privacy loss, could lead to competitive (data) advantages to firms that deploy dark patterns in the engagement-driven digital economy. Similarly, competition law is generally not applicable to collective non-material harms unless it can be shown that such harms would lead to weakening or distortion of competition.

In the context of public enforcement, in contrast to the UCPD and the GDPR, building a credible theory of harm plays an important role in establishing a competition law infringement. Nevertheless, as discussed in section 4.3, competition authorities have yet to bring a case directly involving dark patterns, which could be explained by the practical challenges involved in building such a case, as well as the more prominent role played by other legal regimes (e.g. UCPD) that explicitly target dark pattern practices. Notably, abuse of dominance prohibition is applicable in a very specific set of circumstances i.e. there must be a dominant undertaking engaging in an anticompetitive unilateral practice. Thus, even if dark patterns' harms were to be addressed by competition law, it would not tackle widespread practices employed by smaller market players.

The circumstance that the GDPR, UCPD, DMA, DSA, DA, AI Act all establish national and European authorities with the power to issue (steep) fines also contributes to strengthening the protection of collective interests against collective harms. However, the lack of enforcement of competition law against dark patterns (and, consequently, the harms deriving therefrom) could constitute a missed opportunity to create a well-balanced and complemented system of protection and deterrence against dark patterns' harms.

Finally, a more collective angle in consumer and data protection is also strengthened by Directive 2020/1828 on representative actions for the protection of the collective interests of consumers, enabling entities who have a legitimate interest to pursue action on behalf of consumers for an infringement that harms the collective interests of those consumers.[109] This collective angle is particularly interesting in relation to the collective harms mapped by our taxonomy. Consider, for example, the aforementioned Loss of Trust. If conceptualised as an individual, immaterial harm, Loss of Trust might not meet the requirements and threshold to obtain redress in some Member States. Conversely, Loss of Trust (either in a brand or in the information provided on a product), can reverberate at a collective level, where it can affect the trust of consumers in the market.

5.5. More applicable laws does not mean a better protection for individuals

Dark patterns can violate multiple EU laws at the same time and, ultimately, affect the fairness and balance of powers within the internal market. This means that a single dark

[109]Directive (EU) 2020/1828 of the European Parliament and of the Council of 25 November 2020 on representative actions for the protection of the collective interests of consumers and repealing Directive 2009/22/EC, OJ L 409, 04/12/2020.

pattern type can generate multiple harms, both individual and/or collective, of material and/or non-material nature. It can also breach multiple EU laws at the same time, triggering their respective redress provisions. More applicable laws does not mean a better protection for individuals, but quite the contrary. As mentioned in Section 5.1, the difference in scope, type of harm, terminology, and even in addressees can complicate the concrete enforcement of redress provisions too.

Until the coordination of the enforcement of the different regulations and directives is not clearly defined, the fact that dark patterns can be tackled under multiple EU laws will only complicate the work of competent authorities and national courts, as shown also from the lack of clarity in the interpretation of the scope of Article 25 DSA, in relation to the GDPR and UCPD. This may possibly lead to more leeway to apply the national liability regimes, with less harmonisation and more uncertainty from the EU perspective. A more independent application of national liability regimes is not a bad thing in and of itself, but it is a choice that should be made with awareness, making sure that the highest level of protection and redress is guaranteed to individuals.

5.6. Recommendations for stakeholders

Systematising this volume of sources also illuminates patterns in how empirical researchers and policymakers currently approach the problem of defining and identifying harms caused by dark patterns. Based on our analysis and our systematisation of existing research in this space, we formulate recommendations for regulators, policy makers, and social scientists (economists, behavioural scientists, HCI scholars, legal scholars) working with dark pattern harms.

Regarding regulators and policy makers, a potential identified harm within a complaint can be a heuristic to detect infringements of certain rights (right to privacy, right to data protection).[110] On the other hand, as Esposito et al. propose,[111] drafting a theory of harm, or at least conceptualising the boundaries of harm within a given branch of the law, is the first step for measuring it and this could help enforcement authorities define their priorities. Operationable criteria for measuring the severity of impacts to users has been sought for[112] and could be transposed for dark pattern harms. While it is upon the national courts of the Member States to determine what amounts to compensable damage under Article 82 GDPR or the 11(a) UCPD, we meanwhile question the role of national consumer, competition and data protection authorities (and the authorities that will be appointed under the new EU laws). Even though the actions from supervisory authorities are not linked to awarding compensation to end-users, we contend that evidence of harm can be portrayed by regulators while determining fines. For example, those administrative proceedings could have evidentiary value to support courts in assessing the level of damage[113] in civil proceedings initiated by users that suffered damages from dark patterns. Moreover,

[110]M Hildebrandt, 'The Harm of the Harm Principle in the Context of Risk to Rights', *Linkedln* (2024) https://www.linkedin.com/feed/update/urn:li:activity:7207381953593913344/ accessed 10 September 2024.
[111]Esposito and Sibony (n 19) 5.
[112]G Malgieri and C Santos, 'Assessing the (Severity of) Impacts on Fundamental Rights' (25 June 2024) *SSRN* https://ssrn.com/abstract=4875937 accessed 24 January 2025.
[113]See Article 83 (2)(a) GDPR: 'When deciding whether to impose an administrative fine and deciding on the amount of the administrative fine in each individual case due regard shall be given to the following: (…) the level of damage suffered by them.' A similar provision is included in the UCPD, see section 4.2 above.

supervisory authorities could, in their decision-making processes, explicitly name both dark patterns practices and the specific harms they can cause for deterrence purposes, having in mind the features of dark patterns that make it difficult to provide evidence of dark patterns harms (see beginning of section 4). As noted above, as autonomy is now a protected value, it is required from the legal community to build a theory of harm regarding the recoverability of the Loss of Autonomy, since such loss is now horizontally protected by EU laws (see Annex B).

If the EU legislators think that merely incorporating a redress provision in a given law is enough to protect damaged parties, they are wrong. Provisions such as art. 82 GDPR or 11a UCPD help acknowledging that individuals are entitled to redress, and acknowledge the (more or less) visible harms deriving from certain online practices. Yet these don't resolve the unpredictability and uncertainty of whether certain harms will be recognised, nor whether the costs and effort for the plaintiff will be worth it. Additionally, the new legal provisions dealing with dark patterns do not expressly acknowledge their harms, which as we have shown, are predominantly of moral nature. Currently, the damaged parties face many obstacles that can dissuade them from seeking redress. Elevated costs, complicated national procedures, and a high degree of uncertainty concerning the outcome, dissuade individuals from exercising their right to compensation. Individuals are left on their own, bearing all the burdens of the omnipresent unfair and manipulative practices. Class actions could be a possible solution to this issue, though as a legal institute, they are still in their early stages in the EU and Member States.[114] Class actions are not well harmonised among Member States, and the risk is that the final individual compensation might barely cover their costs. If class actions are to become the solution to dark patterns (and other digital) harms, now it's the time to develop strong principles and guidelines for their application, prioritising individuals that suffered damage. Users can be collectively represented by qualified entities (e.g. noyb or other consumer organisations) and claim for non-material damages related to data protection or consumer-related dark pattern infringements.[115]

Regarding social scientists, we recommend that, when conducting user studies, participants are presented with concrete practices already declared illegal in case law,[116] or at least practices that are most likely in breach of existing EU laws. Since policy makers often refer to HCI and other empirical studies in their reports on dark patterns, if the harms analysed by social scientists are not well aligned with the law, there will still be a missing link between siloed research, and the practice and experience of individuals. To make sure that the harms explored by empirical studies are also enforceable under existing laws, social scientists and legal scholars should work closely, designing experiments that include also, for example, harms acknowledged by the law and legal requisites for its compensation. Rooting the hypothetical harms used during the experiments in legal practice can increase the potential of research to create impact, in the form of supporting data-based policies, and offering tools for enforcement agencies, or even the courts. In addition

[114]Even though the transposition deadline of the EU Directive on Representative Actions (RAD) was 25 December 2022, by the end of November 2024, only 22 out of 27 EU countries had fully implemented it.

[115]Noyb, 'Noyb is Now Qualified to Bring Collective Redress Actions'. https://noyb.eu/en/noyb-now-qualified-bring-collective-redress-actions accessed 24 January 2025; Article 4(2) RAD clarifies that consumer organisations should be considered as entities qualified to represent consumers' collective interests.

[116]An open repository of dark patterns legal cases in the EU/US is available at https://www.deceptive.design/cases accessed 10 September 2024.

to ensuring that user studies align to legal requirements, we also encourage empirical research that explores novel theories of harm, as such studies are instrumental in advancing scientific knowledge and contributing to understanding of potential impacts. Finally, social scientists could develop metrics to measure harm, enabling the systematic collection of evidence for each type of harm. Additionally, they could create tools to assess meaningful platform usage, distinguishing between active and passive engagement. For instance, mindless scrolling or usage may indicate the presence of an attention-capturing dark pattern designed to exploit user engagement.

Behavioural scientists, HCI experts, and legal scholars must work together, to integrate legal perspectives and approaches in the research on harms (as discussed in section 4). Empirical studies should account for the dynamics of dark patterns' harms, including their interrelatedness, dependability, and multifactor influence (as noted in section 3.4). Moreover, it is important to further assess the severity of these harms. Research on dark patterns and their harms needs to be mindful of the need to assure consistency in the terminology deployed to define and characterise them.

6. Conclusions

To support the development of a shared conceptual and linguistic understanding of the harms caused by dark patterns, this article presents our analysis of 39 sources from 12 policy reports and 27 academic studies, and proposes a three-level taxonomy across low-, meso-, and high-level harms. We found 11 low-level harms, and 5 meso-level, distributed across 4 high-level harms. This taxonomy enables a shared vocabulary of the harms caused by dark patterns, clarifies the (mis)alignments between stakeholders and empirical studies, and EU laws scoping dark patterns's harm. We noted divergences, both in the analysed literature review and policy studies, as harms are defined and described with terminological inconsistencies, ultimately hampering the application of national level redress mechanisms. While policymakers and scholarly literature confirm that dark patterns cause harm both at individual and collective level, the effective redress of the harm caused by dark patterns faces several challenges for both material, non-material and collective harms.

Acknowledgements

The authors would like to thank Alessia D'Amico, Martin Brenncke, Václav Janeček, Shu Li, Tina van der Linden, Weiwei Yi, and the anonymous reviewers for their useful comments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Viktorija Morozovaite's work was supported by a Spinoza grant of the Dutch Research Council (NWO), awarded in 2021 to José van Dijck, Professor of Media and Digital Society at Utrecht University.

Appendices

Annex A. List of sources included in the study:

Policy reports

Source ID	Source
PR-0	Organisation for Economic Co-operation and Development (OECD), *Dark Commercial Patterns* (OECD Digital Economy Papers No 336, October 2022)
PR-1	Australian Competition and Consumers Commission (ACCC), *Digital Platform Services Inquiry. Interim Report 6: Report on Social Media Services* (March 2023).
PR-2	Authority for Consumers and Markets (Autoriteit Consument en Markt, ACM), *Guidelines – Protection of the Online Consumer: Boundaries of Online Persuasion* (November 2020)
PR-3	Authority for Consumers and Markets (Autoriteit Consument en Markt, ACM), *EU Fitness Check on Digital Fairness – Protecting Consumers in Digital Environments (Non-paper)* (November 2022)
PR-4	Committee on the Internal Market and Consumer Protection, Kim Van Sparrentak (rapporteur), *Draft Report on Addictive Design of Online Services and Consumer Protection in the EU Single Market (2023/2043(INI))* (November 2023)
PR-5	Consumer and Markets Authority (CMA), *Evidence Review of Online Choice Architecture and Consumer and Competition Harm* (April 2022)
PR-6	Consumer Policy Research Centre (CPRC), *Duped by Design: Manipulative Online Design – Dark Patterns in Australia* (June 2022)
PR-7	European Data Protection Board (EDPB), *Guidelines 3/2022 on Dark Patterns in Social Media Platform Interfaces: How to Recognise and Avoid Them* (13 March 2022).
PR-8	European Union Commission, *Behavioural Study on Unfair Commercial Practices in the Digital Environment: Dark Patterns and Manipulative Personalisation* (April 2022)
PR-9	Federal Trade Commission (FTC), *Bringing Dark Patterns to Light: Staff Report* (September 2022)
PR-10	*Information Commissioner's Office (ICO), Competition and Markets Authority (CMA) and Digital Regulation Cooperation Forum (DRCF), Harmful Design in Digital Markets: How Online Choice Architecture Practices Can Undermine Consumer Choice and Control over Personal Information (Joint Position Paper) (August 2023)*
PR-11	Norwegian Consumers Council (NCC-Forbrukerrådet), *Enough Deception! Norwegian Consumers' Experiences with Deceptive Design* (December 2022)

HCI studies

Source ID	Source
ST-1	Jan M Bauer, Regitze Bergstrøm, and Rune Foss-Madsen, 'Are You Sure You Want a Cookie? – The Effects of Choice Architecture on Users' Decisions about Sharing Private Online Data' in *Proceedings of the Conference on Human Factors in Computing Systems* (2021) 120 *Computers in Human Behavior* 106729 <https://doi.org/10.1016/j.chb.2021.106729> accessed 7 June 2024
ST-2	Aditi M Bhoot, Mayuri A Shinde, and Wricha *P* Mishra, 'Towards the Identification of Dark Patterns: An Analysis Based on End-User Reactions' in *IndiaHCI'20: Proceedings of the 11th Indian Conference on Human–Computer Interaction* (2020)
ST-3	Kerstin Bongard-Blanchy, Arianna Rossi, Salvador Rivas, Sophie Doublet, Vincent Koenig, and Gabriele Lenzini, ''I Am Definitely Manipulated, Even When I Am Aware of It. It's Ridiculous!' – Dark Patterns from the End-User Perspective' in *Designing Interactive Systems Conference 2021* (DIS '21, Virtual Event, USA, 28 June–2 July 2021) ACM, New York, NY, USA <https://doi.org/10.1145/3461778.3462086> accessed 7 June 2024
ST-4	Elise Bonnail, Eric Lecolinet, Wen-Jie Tseng, Samuel Huron, Mark McGill, and Jan Gugenheimer, 'Memory Manipulations in Extended Reality' in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (CHI '23) ACM, New York, NY, USA, Article 875, 1–20 <https://doi.org/10.1145/3544548.3580988> accessed 7 June 2024
ST-5	Christoph Bösch, Benjamin Erb, Frank Kargl, Henning Kopp, and Stefan Pfattheicher, 'Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns' (2016) 4 *Proceedings of Privacy Enhancing Technologies* 237–254 <https://content.sciendo.com/view/journals/popets/2016/4/article-p237.xml> accessed 7 June 2024
ST-6	Gregory Conti and Edward Sobiesk, 'Malicious Interface Design: Exploiting the User' in *WWW '10: Proceedings of the 19th International Conference on World Wide Web* (April 2010) 271–280 <https://doi.org/10.1145/1772690.1772719> accessed 7 June 2024
ST-7	Gregory Day and Abbey Stemler, 'Are Dark Patterns Anticompetitive?' (2020) *Alabama Law Review* 72:1:1, 2–45 <https://www.law.ua.edu/lawreview/files/2020/11/1-DayStemler-1-45.pdf> accessed 7 June 2024

(*Continued*)

Continued.

Source ID	Source
ST-8	Linda Di Geronimo, Larissa Braz, Enrico Fregnan, Fabio Palomba, and Alberto Bacchelli, 'UI Dark Patterns and Where to Find Them: A Study on Mobile Applications and User Perception' in *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (April 2020) 1–14 <https://doi.org/10.1145/3313831.3376600> accessed 7 June 2024
ST-9	Fabiana Di Porto and Albert Egberts, 'The Collective Welfare Dimension of Dark Patterns Regulation' (2023) 29(1-2) *European Law Journal* 114–141 <doi:10.1111/eulj.12478> accessed 7 June 2024
ST-10	Ted Harris, Simon Nilsson, and Talia Beck, 'Dark Patterns in Online Shopping: Do They Work and Can Nudges Help Mitigate Impulse Buying?' (2022) *Behavioral Public Policy* 1
ST-11	Kawon (Kathy) Kim, Woo Gon Kim, and Minwoo Lee, 'Impact of Dark Patterns on Consumers' Perceived Fairness and Attitude: Moderating Effects of Types of Dark Patterns, Social Proof, and Moral Identity' (2023) *Tourism Management* 98, 104763 <https://doi.org/10.1016/j.tourman.2023.104763> accessed 7 June 2024
ST-12	Monica Kowalcyzk, Johanna T. Gunawan, David Choffnes, Daniel J Dubois, Woodrow Hartzog, and Christo Wilson, 'Understanding Dark Patterns in Home IoT Devices' in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (CHI '23, Hamburg, Germany, 23–28 April 2023) ACM, New York, NY, USA, 27 pages <https://doi.org/10.1145/3544548.3581432> accessed 7 June 2024
ST-13	Veronika Krauß, Pejman Saeghe, Alexander Boden, Mohamed Khamis, Mark McGill, Jan Gugenheimer, and Michael Nebeling, 'What Makes XR Dark? Examining Emerging Dark Patterns in Augmented and Virtual Reality through Expert Co-Design' (2024) 31 ACM Trans. Comput.-Hum. Interact. 32 <https://doi.org/10.1145/3660340>accessed 7 June 2024
ST-14	Ida Borberg, Rene Hougaard, Willard Rafnsson, and Oksana Kulyk, ''So I Sold My Soul': Effects of Dark Patterns in Cookie Notices on End-User Behavior and Perceptions' in *Proceedings of the 2022 Symposium on Usable Security* (2022)
ST-15	Lin Kyi, Sushil Shivakumar, Cristiana Santos, Franziska Roesner, Frederike Zufall, and Asia J. Biega, 'Investigating Deceptive Design in GDPR's Legitimate Interest' in *CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (April 2023) Article No. 583, 1–16 <https://doi.org/10.1145/3544548.3580637>accessed 7 June 2024
ST-16	Chris Lewis, 'Monetary Dark Patterns' in *Irresistible Apps* (Apress, Berkeley, CA, 2014) <https://doi.org/10.1007/978-1-4302-6422-4_10> accessed 7 June 2024
ST-17	Jamie Luguri and Lior Strahilevitz, 'Shining a Light on Dark Patterns' (2021) 13 *Journal of Legal Analysis* 43, University of Chicago Coase-Sandor Institute for Law & Economics Research Paper No 879, U of Chicago, Public Law Working Paper No 719 <https://ssrn.com/abstract=3431205> accessed 7 June 2024
ST-18	Maximillian Maier and Rikard Harr, 'Dark Design Patterns: An End-User Perspective' (2020) 16 *Human Technology* 170–199
ST-19	Arunesh Mathur, Jonathan Mayer, and Mihir Kshirsagar, 'What Makes a Dark Pattern … Dark?: Design Attributes, Normative Considerations, and Measurement Methods' in *CHI Conference on Human Factors in Computing Systems* (CHI '21, Yokohama, Japan, 8–13 May 2021) ACM, New York, NY, USA, 27 pages<https://doi.org/10.1145/3411764.3445610> accessed 7 June 2024
ST-20	Thomas Mildner and Gian-Luca Savino, 'Ethical User Interfaces: Exploring the Effects of Dark Patterns on Facebook' in *CHI Conference on Human Factors in Computing Systems* Extended Abstracts (CHI '21 Extended Abstracts, Yokohama, Japan, May 8–13, 2021) ACM, New York, NY, USA, 7 pages <https://doi.org/10.1145/3411763.3451659> accessed 7 June 2024
ST-21	Thomas Mildner Gian-Luca Savino, Philip R. Doyle, Benjamin R. Cowan, Rainer Malaka, 'About Engaging and Governing Strategies: A Thematic Analysis of Dark Patterns in Social Networking Services' in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (CHI '23, Hamburg, Germany, April 23–28, 2023) ACM, New York, NY, USA, 15 pages <https://doi.org/10.1145/3544548.3580695>accessed 7 June 2024
ST-22	Monge Roffarello A, Lukoff K, and de Russis L, 'Defining and Identifying Attention Capture Deceptive Designs in Digital Interfaces' in *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2023) Article 194, 1–19 <https://doi.org/10.1145/3544548.3580729> accessed 7 June 2024
ST-23	Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal, 'Dark Patterns after the GDPR: Scraping Consent Pop-Ups and Demonstrating Their Influence' in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI '20) Association for Computing Machinery, New York, NY, USA, 1–13 <https://doi.org/10.1145/3313831.3376321> accessed 7 June 2024
ST-24	Lorena Sánchez Chamorro, Carine Lallemand, and Colin M Gray, '"My Mother Told Me These Things Are Always Fake" – Understanding Teenagers' Experiences with Manipulative Designs' in *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (2024)
ST-25	Christian Voigt, Stephan Schlögl, and Aleksander Groth, 'Dark Patterns in Online Shopping: Of Sneaky Tricks, Perceived Annoyance and Respective Brand Trust' in *HCI in Business, Government and Organizations: 8th International Conference, HCIBGO 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, 24–29 July 2021, Proceedings* (Springer-Verlag, Berlin, Heidelberg 2021) 143–155 <https://doi.org/10.1007/978-3-030-77750-0_10> accessed 7 June 2024

(*Continued*)

Continued.

Source ID	Source
ST-26	Rahul De', Souren Paul, Suprateek Sarker, Virpi Kristiina Tuunainen, Walter D Fernández, and Joe Nandhakumar (eds), *Proceedings of the 44th International Conference on Information Systems, ICIS 2023: Rising Like a Phoenix: Emerging from the Pandemic and Reshaping Human Endeavors with Digital Technologies* (Hyderabad, India, 10–13 December 2023) Association for Information Systems, 2023
ST-27	Amit Zac, Yun-Chun Huang, Amédée von Moltke, Christopher Decker, Ariel Ezrachi, 'Dark Patterns and Online Consumer Vulnerability' (22 August 2023) <https://ssrn.com/abstract=4547964> accessed 7 June 2024

Annex B. Overview of harms-related terminology, type and scope in current EU laws

Laws	Harm-related terminology from legal provisions and recitals	Type of harm	Scope of harm	Subjects harmed
DMA	R(70):impair user autonomy, decision-making, or choice Art.13(6) gatekeepers should not '… or make exercise of those rights or choices unduly difficult'	Autonomy loss; Economic loss; Impossibility to exercise rights	Collective	End users and business users
DSA	R(67): unwanted behaviours; into undesired decisions; not be in the recipients' interests; which have negative consequences for them; autonomous and informed choices or decisions; Art. 25: distorts or impairs the ability of the recipients of their service to make free and informed decisions.	Negative consequences; Autonomy loss	Collective	Consumers and Business users
AIA	Art. 5(1)(a)-(b): causing the person to take a decision that that person would not have otherwise taken; causes or is likely to cause that person, another person or group of persons significant harm; impairing their ability to make an informed decision	Significant harm; Autonomy loss	Individual and collective	Users
DA	Arts 4(4), (6)(1)(a): shall not make the exercise of choices or rights under this Article by the user unduly difficult, including by offering choices to the user in a non-neutral manner or by (coercing) subverting or impairing the autonomy, decision-making or choices of the user (…). Art. 38(1): (…) natural and legal persons shall have the right to lodge a complaint, individually or, where relevant, collectively, (…) if they consider that their rights under this Regulation have been infringed. R(38): (…) third parties or data holders should not rely on so-called 'dark patterns' in	Impossibility to exercise rights; Autonomy loss	Individual	Users, i.e. Natural or legal persons that own or are entitled by the law to use a connected product or related service

(*Continued*)

Continued.

Laws	Harm-related terminology from legal provisions and recitals	Type of harm	Scope of harm	Subjects harmed
	designing their digital interfaces (…). Those manipulative techniques can be used to persuade users, in particular vulnerable consumers, to engage in unwanted behaviour, to deceive users by nudging them into decisions on data disclosure transactions or to unreasonably bias the decision-making of the users of the service in such a way as to subvert or impair their autonomy, decision-making and choice.			
GDPR	R(75) The risk to the rights and freedoms of persons, of varying likelihood and severity, may result from personal data processing which could lead to physical, material or non-material damage; R(85) A personal data breach may, (…) result in physical, material or non-material damage to natural persons; Art. 82: any person who has suffered material or non-material damage	Damage (physical, material or non-material damage)	Individual	Data subject
UCPD	Art. 11a: '*1*. Consumers harmed by unfair commercial practices, shall have access to proportionate and effective remedies, including compensation for damage suffered by the consumer and, where relevant, a price reduction or the termination of the contract.' R(10): This Directive consequently complements the Community *acquis*, which is applicable to commercial practices harming consumers' economic interests.	Material distortion of economic behaviour; Cause the consumer to take a transactional decision that he would not have taken otherwise	Individual	Consumers (average or vulnerable)

Annex C. The taxonomy of dark patterns