Defining and Identifying Attention Capture Deceptive Designs in Digital Interfaces

Alberto Monge Roffarello
Politecnico di Torino
Torino, TO, Italy
alberto.monge@polito.it

Kai Lukoff
Santa Clara University
Santa Clara, CA, USA
klukoff@scu.edu

Luigi De Russis
Politecnico di Torino
Torino, TO, Italy
luigi.derussis@polito.it

ABSTRACT

Many tech companies exploit psychological vulnerabilities to design digital interfaces that maximize the frequency and duration of user visits. Consequently, users often report feeling dissatisfied with time spent on such services. Prior work has developed typologies of damaging design patterns (or dark patterns) that contribute to financial and privacy harms, which has helped designers to resist these patterns and policymakers to regulate them. However, we are missing a collection of similar problematic patterns that lead to *attentional harms*. To close this gap, we conducted a systematic literature review for what we call 'attention capture damaging patterns' (ACDPs). We analyzed 43 papers to identify their characteristics, the psychological vulnerabilities they exploit, and their impact on digital wellbeing. We propose a definition of ACDPs and identify eleven common types, from Time Fog to Infinite Scroll. Our typology offers technologists and policymakers a common reference to advocate, design, and regulate against attentional harms.

CCS CONCEPTS

• **Human-centered computing** → **HCI theory, concepts and models**; **Empirical studies in HCI**; Collaborative and social computing design and evaluation methods; • **Information systems** → *Web interfaces.*

KEYWORDS

digital wellbeing, damaging patterns, deceptive design, dark patterns, technology overuse, attention

1 INTRODUCTION

Nowadays, many users feel conflicted about the amount of time they passively spend on their devices [53, 66, 126], and terms like "overuse [66]," "compulsive use [119]," and even "tech-addiction[1] [62]" are often used to highlight the negative impacts that technology can have on people's mental health [63] and social interactions [120]. Contemporary smart devices and online platforms like social networks, in particular, can easily become a source of distraction, e.g., due to notifications [95, 97] or self-interruptions [94]: this can interfere with people' daily activities like studying, working, and driving [4, 48], and makes them less productive [79] and more stressed [80]. The aforementioned studies, as well as a growing public discussion on mainstream media [25, 42], have motivated and inspired HCI studies on the intentional "non-use" of technology [106] and digital wellbeing [21].

Despite this growing interest in digital wellbeing, tech companies still continue to design experiences that maximize attention capture and contribute to these problems. Attention, i.e., a state in which the individual's cognitive resources are selectively focused on some perceived stimuli coming from the environment [8], is one of the most valuable resources of the digital age. In the attention economy [32], businesses transform attention into a currency, whereby users "pay" for a service with the time they spend on it [54]. Researchers point out that this attention capture is *by design* [11], exploiting designs that diminish people' sense of agency [12, 73] and control over technology use [27]. Such harmful designs have become known as "dark patterns", i.e., recurring interaction design patterns in digital user interfaces that designers use to intentionally manipulate users into performing actions that go against their best interests [46]. However, the association of "dark" with harm is problematic, so to use more inclusive language, designers have introduced the terms "deceptive design (patterns)" [19] and "damaging patterns" [108], the latter of which we adopt here (see the Discussion section of this paper for a full reflection on our choice of terminology). The concept of damaging patterns was first introduced by the design practitioner Harry Brignull in 2010, who published a gallery of them on the *www.deceptive.design* website [19] (formerly *www.darkpatterns.org*) and uses a Twitter hashtag campaign to highlight examples: formerly #darkpatterns and now #deceptivedesign.

The concept of damaging patterns has since attracted growing interest from practitioners, researchers, and regulators [72]. The name-and-shame approach of the original website and Twitter campaign has led to increased awareness among designers [46], gatherings of researchers across fields (e.g., [83]), and widespread media attention (e.g., [88, 102, 103]). It has also motivated regulatory hearings across the world (e.g., by the US Federal Trade Commission [40] and the European Data Protection Board [37])

---

[1]The usage of an addiction framing for everyday behaviors like PC and smartphone usage is currently debated (see the work of Lanette et al. [64]).

and led to regulations against certain damaging patterns [1], with others under consideration [117].

In a review of the past, present, and future of damaging patterns, Narayanan et al. write *"dark patterns enable designers to extract three main resources from users: money, data, and attention"* [89] (p. 12). Thus far, researchers [46, 82] and practitioners [19] have focused on the damaging patterns that lead to financial and privacy harms. For example, the original patterns proposed by Brignull include *Sneak into Basket* [19, 46], when e-commerce websites surreptitiously add items into the shopping cart that the user never selected, and *Privacy Zuckering* [19, 46], user interfaces that trick users into (unintentionally) sharing private information. In other notable examples that examine damaging patterns within a domain, Mathur et al. crawled over 10,000 shopping websites to identify the prevalence of financial deceptive designs and Bösch et al. identify a collection of privacy deceptive designs [17]. Design practitioners and researchers have made considerable progress towards identifying the manipulative design patterns that extract money and data from users.

However, the *attentional harms* of damaging patterns have been largely neglected. Of Brignull's original 12 patterns, only one (Disguised Ads) focuses on interfaces that trick users into spending time and attention in ways that go against their best interest. One notable exception is the work of Lewis and colleagues that identifies patterns that they call "temporal dark patterns" in the context of game design (e.g., Playing by Appointment) [67, 130]. Yet there is a lack of a systematic overview of the designs that lead to attentional harms across domains, including social media, which has been implicated in much research on problematic (over)use [58]. This has led to calls for the design community to identify the patterns that manipulate users into spending their time and attention in ways that go against their interests [73].

To close this gap, this paper conducts a systematic literature review that develops and defines the concept of **Attention Capture Damaging Patterns (ACDP)** within papers at the intersection of damaging design patterns and digital wellbeing. Drawing upon the findings of previous work, we first present a *definition* for ACDPs. We then explore the characteristics of these patterns by proposing a set of five criteria that shed light on how tech companies implement these kind of manipulative interfaces and describe how ACDPs exploit users' psychological vulnerabilities. Finally, we extract a *typology* of 11 ACDPs, from Time Fog to Infinite Scroll, by describing how they apply to different digital services and interfaces.

Overall, our work defines ACDPs 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. Some of these patterns, e.g., Disguised Ads and Recommendations, use deception to trick the user into false beliefs. Others, e.g., Guilty Pleasure Recommendations, seduce the user into temptations that they wish to avoid.

Although some operate by deception and others seduction, ACDPs all exploit the fact that many of the biases and heuristics of an individual are predictable to manipulate the user into impulsive short-term decisions that go against their long-term aspirations for how they wish to spend their time. This potentially triggers negative behaviors like compulsive usage patterns and technology overuse, leads the user to feel a sense of regret, and negatively influences psychological experience and self-beliefs, from sense of agency to self-efficacy.

Developing a *"lingua franca of attention capture design patterns"* [73] has several pathways for impact, from motivating stakeholders to find alternatives to the contemporary attention economy to the development of better tools for digital self-control. We discuss such opportunities in the Discussion section of this paper. Furthermore, we created the *http://attentioncapture.com/* website, an online gallery of the 11 ACDPs described in this paper, to increase the reach of our work among the public and design professionals.

2 METHODS

We conducted a systematic literature review to develop a definition of attention capture damaging patterns, identify types of design patterns that meet our criteria, and share examples of each. To identify and select relevant papers, we followed the PRISMA literature review guidelines [69, 86] (see Figure 1 for our procedural flowchart).

2.0.1 Paper Selection. As in recent literature reviews in the HCI domain (e.g., [118]), we identified relevant papers by searching the electronic database of the Association for Computing Machinery (ACM) Guide to the Computing Literature[2], the most comprehensive bibliographic source collection in the field of computing and HCI research. It integrates the traditional ACM Digital Library with conference proceedings, journals, magazines, books, and abstracts of key publishers like IEEE, Springer, and Elsevier.

All the searches included manuscripts published from January 1, 2000 to June 13, 2022 whose "content type" was "Research Article." We defined our search terms (from "dark patterns" to "digital overload", Table 1) based on the terminology used in studies about deceptive designs and digital wellbeing, respectively. For deceptive designs, in particular, we took inspiration from the search terms adopted by Mathur et al. [82]. For digital wellbeing, instead, we looked at words in titles, abstracts, and keywords in the papers reviewed by Lyngs et al. [76]. We believe the adopted terms capture most of the related work at the intersection of deceptive designs and digital wellbeing around computer science. Overall, the initial search identified a total of 1,334 records. We firstly analyzed the retrieved collection by removing 414 duplicates. Other 822 records were excluded through a screening of the titles and the abstracts. Such a screening was performed by applying a set of inclusion/exclusion criteria. Papers were eligible for inclusion if they met *all* of these criteria:

- papers that focus, directly or indirectly, on deceptive designs and dark patterns, i.e., user interfaces that intentionally manipulate users into performing actions that go against their best interests;
- papers that include at least some considerations on people's digital wellbeing, with a particular focus on the technology overuse aspect;
- papers that discuss the relationship between specific digital services characteristics, e.g., designs and system functionality, and people's digital wellbeing.

---

[2]https://libraries.acm.org/digital-library/acm-guide-to-computing-literature, last visited on August 19,2022.
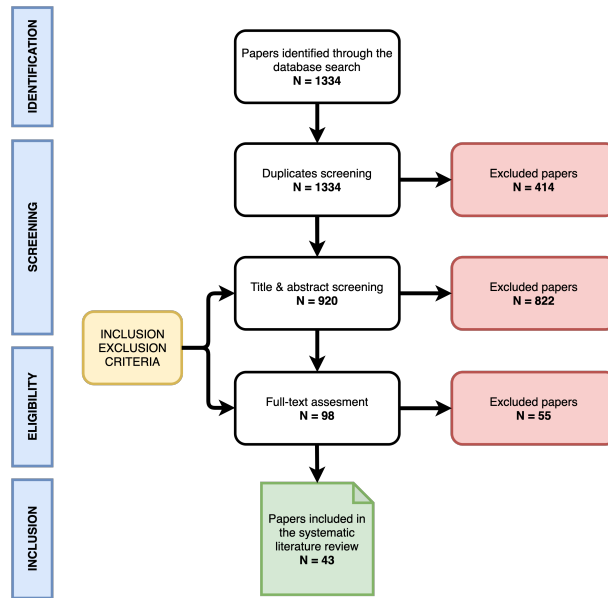
Figure 1: Procedural flowchart following the PRISMA guidelines.

Table 1: The search queries used to search the electronic database of the ACM Guide to the Computing Literature. All the searches included manuscripts published from January 1, 2000 to June 13, 2022 whose "content type" was "Research Article."

Search Query	# Results
"internet addiction" OR "smartphone addiction" OR "social media addiction" OR "technology addiction" OR "app addiction"	691
"dark pattern*"	170
"attention economy" OR "attention-capture"	159
"smartphone overload" OR "smartphone overuse" OR "phone overload" OR "phone overuse"	73
"compulsive behaviour" OR "compulsive behavior"	72
"digital overload" OR "digital overuse" OR "technology overload" OR "technology overuse"	57
"unethical design*" OR "evil design*" OR "manipulative design*"	35
"internet overload" OR "internet overuse"	32
"digital distraction*"	32
"unethical interface*" OR "evil interface*" OR "manipulative interface*"	13
"social media overload" OR "social media overuse" OR "social networks overload" OR "social networks overuse"	10

Papers were excluded if they matched *one or more* of these exclusion criteria:

- papers about deceptive designs that do not address problematic exploitation of time and attention, e.g., papers that focus instead on privacy harms, such as [112, 127];
- papers that describe interventions that support digital well-being, e.g., timers, but do not analyze problematic design patterns in existing apps, such as [59, 60];
- papers that focus on users, only, e.g., those that analyze people's problematic Internet use without explicitly linking problems to digital interfaces [56, 100].

At the end of the screening, we further analyzed whether the remaining 98 publications met the defined inclusion and exclusion criteria. This full-text read-through led us to remove a further 55

publications. This left us with a final set of 43 research publications[3] that were included in the systematic literature review: 30 conference papers (27 full papers and 3 extended abstracts) and 13 journal articles. We checked all the conference and journal websites to ensure that all of the included publications, including extended abstracts, had been peer-reviewed. The conferences with the most publications in our corpus were CHI[4] (17) and DIS[5] (3), while all other conferences, from WWW[6] to MobileHCI[7], contributed only a single paper. Journal articles came mostly from PACM[8] (4) and

---

[3]The included publications are highlighted in the References list through a check mark (✔).
[4]https://dl.acm.org/conference/chi, last visited on August 19,2022
[5]https://dl.acm.org/conference/dis, last visited on August 19,2022
[6]https://dl.acm.org/conference/www, last visited on August 19,2022
[7]https://dl.acm.org/conference/mobilehci, last visited on August 19,2022
[8]https://www.acm.org/publications/pacm, last visited on August 19,2022

Computers in Human Behavior[9] (4). As reported in Figure 2, our final corpus shows that the intersection between deceptive designs and digital wellbeing has emerged as a research area in the last ten years, with growing interest in the last five years.
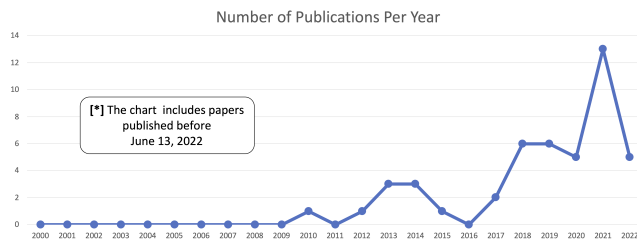

Figure 2: The number of publications that fulfills our inclusion/exclusion criteria per year. The graph highlights that the intersection between deceptive designs and digital wellbeing is a relatively recent research area.

2.0.2 Data Analysis & Coding Process. To systematically analyze our corpus, we created a spreadsheet to code different aspects extracted from the analyzed papers. We used the first columns of the sheet to characterize papers by their authors, title, abstract, publication type and year, and to summarize the presented contributions. We extracted relevant information on the *characteristics* of attention capture damaging patterns they proposed, and we recorded information about the *mechanisms* adopted by tech companies to design "addictive" digital services, i.e., which techniques and users' psychological vulnerabilities are exploited. Furthermore, we also extracted information about the *impacts* of ACDPs may have on users' digital wellbeing. The described extraction sheet template was created by one of the authors by coding ten randomly selected papers. The sheet was then checked by the other authors, who implemented some minor adjustments. Each paper of the corpus was finally analyzed using the final version of the extraction sheet template. After developing a clear picture of what constitutes an attention capture damaging pattern, we performed another pass on the 43 papers of our corpus to extract examples of ACDPs, with the aim of producing a comprehensive typology of this specific kind of damaging design patterns.

3 SYSTEMATIC LITERATURE REVIEW OF ATTENTION CAPTURE DECEPTIVE DESIGNS

In this section, we present the results of our systematic literature review. Findings are organized around the two main contributions of this work, i.e., *definition* of attention capture damaging patterns (Section 3.1) and *typology* of the exploited design patterns (Section 3.2). Table 2 summarizes the review by linking definition criteria, patterns, and papers from which we extracted them.

3.1 Attention Capture Damaging Patterns: Definition and Criteria

As pointed out by Gray et al. [46], the original definition of damaging patterns, i.e., functionality that exploits people's psychological vulnerabilities to promote choices that are not in the user's best interest, leaves many questions unanswered, e.g., "*what is the user being 'tricked' into doing, and with what motivation*" (p. 3). In our work, we refer to those patterns through which designers explicitly assert control over the user's experience [44] to keep users as customers of the service [10, 81] and generate more income [85, 128]. Widdicks et al. [128], for example, reported a statement of a former Facebook employee (taken from [5]), who said: "*you have a business model designed to engage you and get you to basically suck as much time out of your life as possible and then selling that attention to advertisers.*"

We define an attention capture damaging pattern (ACDP) as:

> *A recurring pattern in digital interfaces that a designer uses to exploit psychological vulnerabilities and capture attention, often leading the user to lose track of their goals, lose their sense of time and control, and later feel regret.*

The goal of ACDPs is to maximize continuous usage [20, 27, 34, 73], daily visits [20, 73], and interactions [20, 44, 73, 131] (e.g., clicks, shares, likes, etc.). They make users "*more likely to visit [a digital service] again and click on similar types of rewarding content*" [20], thus creating a "*trap for the user that enables the stakeholder's goal*" [45].

Table 3 reports five criteria that can be used to further characterize and identify ACDPs[10]. The first two criteria (C1, C2) are related to the *mechanisms* exploited by ACDPs:

C1 - Exploit Psychological Vulnerabilities. In the same way that nudges leverage psychological heuristics and biases to guide people toward actions that are in their best interests (e.g., eating healthier) [23], attention capture damaging patterns exploit this same psychology to induce actions that go against their best interests (e.g., spending more time in an app than they would like). By preying on the fact that many user biases are predictable [7], these designs shove people towards actions that users may not choose if they were making a considered decision [20]. Notasil and Payne [91], for example, concluded that an emotional memory bias might increase the attractiveness of the newsfeed. Lukoff et al. [73] proposed that recommendations on YouTube might exploit short-term bias, wherein people favor the choice that offers immediate gratification, e.g., watching a new catchy video, at the expense of long-term goals. Similarly, Bongard-Blanchy et al. [16] stated that ACDPs seduce users with benefits like ease of use and immediate gratification.

Also, ACDPs often leverage a variable schedule of rewards [13, 20]. According to Skinner's operant conditioning theory [113], the most effective way of reinforcing behavior is to follow a variable schedule of rewards: even the task of predicting an outcome is *itself* rewarding and triggers the release of dopamine. Burr et al. [20] reported that exposure to variable reward might occur every time

---

Table 2: A summary of our systematic literature review encompassing definition criteria (from C1 to C5, see Table 3 for further details) and attention capture deceptive designs (from P1 to P11, see Table 4 for further details).

	Definition Criteria					Attention Capture Deceptive Designs										
	Exploit psychology	Automatize experience	Divert attention	Lost time and control	Sense of regret	Infinite scroll	Casino pull-to-refresh	Neverending autoplay	Guilty pleasure recs	Disguised ads and recs	Recapture notifications	Playing by appointment	Grinding	Attentional roach motel	Time fog	Fake social notifications
Baroni et al. [10]		✓												✓		✓
Burr et al. [20]	✓	✓	✓	✓			✓		✓	✓						
Nontasil et al. [91]	✓		✓			✓	✓		✓							
Conti et al. [31]	✓		✓													
Mathur et al. [81]																
Gray et al. [46]	✓												✓			✓
Di Geronimo et al. [33]	✓									✓			✓	✓		
Gunawan et al. [47]													✓	✓		✓
Bhoot et al. [78]	✓													✓		✓
Mathur et al. [82]	✓	✓	✓			✓						✓	✓			
Mildner et al. [85]			✓	✓		✓				✓				✓		
Bongard-Blanchy et al. [16]	✓	✓	✓													
Gray et al. [44]		✓								✓						
Kollnig et al. [61]	✓		✓					✓								✓
Gray et al. [45]	✓		✓													
Zeng et al. [131]			✓							✓						
Susser et al. [114]	✓	✓							✓	✓						
Fitton et al. [41]										✓		✓	✓			✓
Widdicks et al. [128]	✓	✓		✓		✓		✓				✓	✓			
Lukoff et al. [73]	✓	✓	✓	✓	✓	✓			✓	✓	✓		✓			
Tran et al. [119]	✓			✓	✓	✓	✓									
Pinder [98]										✓						
Cho et al. [29]	✓			✓	✓	✓	✓	✓	✓	✓						
Lee et al. [66]	✓						✓									
Monge Roffarello et al. [104]						✓										
Park et al. [96]						✓			✓							
Aranda et al. [6]	✓			✓	✓	✓			✓			✓				
Diefenbach et al. [34]	✓						✓									
Cheng et al. [27]			✓			✓		✓								
Bedjaoui et al. [13]	✓	✓	✓			✓			✓			✓			✓	
Hasan et al. [50]		✓	✓					✓	✓							
Lyngs et al. [77]	✓		✓	✓		✓			✓	✓		✓				
Lukoff et al. [75]		✓				✓										
Kim et al. [57]	✓															
Harwood et al. [49]	✓															
Zagal et al. [130]													✓	✓		✓
Meshi et al. [84]	✓															
Urmanov and Hoyoung [125]			✓							✓	✓					✓
Baughan et al. [11]	✓	✓	✓	✓	✓	✓					✓					
Farivar et al. [39]	✓		✓							✓						
Hung et al. [52]	✓	✓	✓													
Chaudhary et al. [26]		✓	✓	✓	✓				✓						✓	
Zhang et al. [132]		✓	✓			✓			✓	✓	✓					✓
# of papers	26	14	16	13	8	16	6	5	12	14	7	5	7	6	1	9

a user engages with an intelligent system agent, as the user typically does not know what items will be presented. Some analyzed papers [6, 20, 34, 91, 128] even relate attention capture damaging patterns to slot machines, saying, for example, that the newsfeed of popular social networks exploit the same psychological vulnerabilities targeted in gambling addictions: users do not know in advance the posts that will be displayed, and each visualized post may be rewarding or not, e.g., a photo by a friend *vs.* an unwanted advertisement. This uncertainty fosters the temptation to constantly

Table 3: A set of 5 criteria characterizing attention capture damaging patterns.

Criteria	Category	Description
C1	*Mechanisms*	Exploit known psychological biases and heuristics.
C2	*Mechanisms*	Automate the entire user experience.
C3	*Impacts*	Lead users to lose track of their goals.
C4	*Impacts*	Lead to a lost sense of time and control.
C5	*Impacts*	Lead to a sense of regret about the time spent on a digital service.

check [34] and leads to continuous use. In turn, this even leads to reward depletion [29, 66]: where users find themselves scrolling through posts and videos that they have already seen, while they hope for new items to appear.

C2 - Automate the User Experience. Attention capture damaging patterns may automate the user experience to induce meaningless normative dissociation experiences that direct the behavior and keep users on the platform. Normative dissociation is a phenomenon in which a person temporarily experiences a disconnection from physical and emotional experiences. During normative dissociation, people experience a loss of self-awareness and reflection, and they are less inclined to exercise intentional choice. These experiences are typically only realized in hindsight, i.e., once self-reflection is reengaged [22]. Normative dissociation can characterize different mental states, including daydreaming, flow states, and becoming absorbed in watching a movie [22].

While these experiences may sometimes be beneficial for the user, e.g., in the case of flow, Baughan et al. [11] warn that designers may intentionally adopt patterns that promote "zone states," i.e., absorption in personally meaningless activities with little to no intrinsic value. In the context of a study on Twitter, the authors reported that participants described feelings of being absorbed in a "zombie"-like state when passively scrolling the newsfeed. Similarly, "The 30-Minute Ick Factor" reported by Tran et al. [119] describes the negative feelings that users experience after noticing that they have spent a notable amount of time on social media unconsciously.

To induce meaningless normative dissociation, ACDPs often remove the need for autonomous decision making [20, 44], by promoting "endless" sessions [128]. Gray et al. [44], for example, report that manipulative designs, including those that can lead to attentional harms, adopt mechanisms that *"automate the process of performing essential tasks without the user' confirmation"* (p. 67). As reported by Chaudhary et al. [26], there is a need to discuss *"the close correlation between ease of usability and dark persuasive patterns"* (p. 788). In analyzing deceptive designs on video streaming platforms, the authors found several functional and helpful features that, in reality, may evolve into damaging patterns with adverse consequences on users' digital wellbeing. The negative side of ACDPs is that user interface improvements and simplifications are sometimes a deliberate choice of designers and tech companies to promote a frequent and continuous use of technology [13, 16]. Designers, in particular, often try to improve their services' design without thinking about their choices' unintended adverse consequences [128]. Consequently, Chaudhary et al. [26] warn that the trade-off between usability and persuasion is critical,

especially when there are ambiguities in the designer's intentions. After prolonged use, in particular, features like content autoplay may become *"habit-forming designs"* [26].

The three remaining criteria (C3, C4, C5) address the *impacts* that ACDPs may have on users' digital wellbeing:

C3 - Lose Track of Goals for Use. Attention capture damaging patterns lead users to lose track of their goals by demanding their attention and introducing frequent distractions [31, 77]. As with other damaging patterns, ACDPs may *attract* [131] or *divert* [13] attention. As a result, users experience situations in which they are tricked into taking actions that are aligned with the stakeholders' goals rather than their own [13, 31, 61, 77]. Lyngs et al. [77] found that a typical newsfeed on Facebook contains much attention-grabbing and distracting content. More generally, Conti et al. [31] classified different malicious designs and strategies that may cause distraction, e.g., catchy videos and animations in advertisements. Frequent distractions are in turn correlated with a decrease in users' *productivity* [27, 73].

C4 - Lost Sense of Time and Control. Attention capture damaging patterns make a person experience a lost sense of time and control. A participant in the study of Cho et al. [29] explains such a feeling in this way:

> I keep pressing next and flipping a story to another. I just keep pressing... to just waste time rather than actually viewing it (p. 12).

Damaging designs that lead to attentional harm, in particular, negatively influence users' sense of agency [73]. User agency or self-agency is defined as a person's self-perception of being the initiator of its actions [115]. ACDPs may present information in a way that reduces user autonomy of choice by adopting coercive and deceptive strategies [10, 20, 114]. Lukoff et al. [73] point out that low sense of agency over technology use is, in turn, associated with negative experiences and a general sense of dissatisfaction over social media use. By surveying and interviewing YouTube users, the authors found that features like recommendations and autoplay often make users feel less in control as they undermine their sense of agency, e.g., because suggestions of new videos are typically "hard to decline."

C5 - Sense of Regret. The exposure to an attention capture damaging pattern is typically associated with a later sense of regret, e.g., about the time spent on a digital service or a specific interaction with it. As explained by a participant of the study of Tran et al. [119] speaking about Instagram usage, for instance:

"[It] gave me like, temporary satisfaction. Like, 'Oh yeah, all these people like my photo,' or 'All these people think my story is funny.' And yeah, it's great in that moment, but then after it dies down, you're just kind of just like, 'What's the point?'" (p. 8).

As reported by Cho et al. [29], regret happens when *"the rewards of a taken action are outweighed by the expected rewards of what could have happened alternatively"* (p. 456:2). The regret theory [105], in particular, defines regret as a counterfactual feeling that *"the past might have unfolded differently, particularly if a different decision had been made"* (p. 2). Being exposed to ACDPs increases the chances of using (or continuing to use) a digital service at times when users would not have otherwise [27], and this causes regret, e.g., when users spend more time than they planned [6]. Indeed, websites and mobile apps on which we spend the most time, e.g., social networks, are also those we regret using the most [20]. This tendency is confirmed by the study of Cho et al. [29], which investigates the relationship between different features of social media and regret. For example, the authors found that repeated use of "following-based" features like newsfeed and stories quickly deplete content and cause regret. Similarly, "recommendations-based" features with bite-sized contents, e.g., Facebookâ€™s Watch Videos, induce users to use the service "just a bit more," promoting a behavioral cycle that makes users experience a later sense of regret.

3.2 A Typology of Attention Capture Damaging Patterns

Our next step was to develop a typology of ACDPs based on the literature we reviewed. Although the terms typology and taxonomy are often used interchangeably [35], we purposefully chose *typology* to emphasize that our patterns are "ideal types," as in types that represent elements common to most cases across the literature. Unlike a taxonomy, there are no strict decision rules to determine whether a given design pattern fits type A or type B.

In developing the typology, we faced three significant methodological decisions. First, we considered whether to name patterns in academic language or everyday language. Here, we drew upon the early work of the architect Christopher Alexander who advocated for patterns that are 'alive,' which spark inspiration for the designers and capture the imagination of the public [2]. Certainly Brignull could have given "Sneak into the Basket" a more technical name, e.g., "opt-out e-commerce," however we doubt that it would have reached as wide of an audience and served its function as a common reference for the design community and beyond. We thus chose to give patterns evocative names using everyday language.

Second, we had to decide whether to include design patterns that met only one part of our definition. For example, in their review of shopping websites, Mathur et al. describe "Countdown Timer" and "Limited-time Message" as deceptive designs that use a sense of time urgency to capture attention [81]. While these patterns do leverage sense of time as a *mechanism*, we decided to exclude them from our typology as their impact is primarily a financial harm. Instead, we focused on patterns where the impact is also an attentional harm.

Finally, we needed to determine how much context to include in our patterns. A universal challenge for damaging patterns is that

not all patterns are harmful all of the time. For instance, Brignull describes how an interface element like "opt-out defaults" (a checkbox or radio button that is pre-selected for the user) might be ethical in one context, but not in another [18]. On a form for organ donation, it might be ethical to set "donor" as the default. However, the same interface element might be unethical if used to automatically add an iPad case to a user's shopping cart when they purchase an iPad. Thus, instead of "default settings" Brignull formulates the pattern as "Sneak into basket" that is specific to the e-commerce shopping experience: *"You attempt to purchase something, but somewhere in the purchasing journey the site sneaks an additional item into your basket."* Similarly, regulation by the EU consumer protection agency forbids opt-out for shopping baskets and email newsletters, rather than as a design pattern in general [1]. The same pattern can have different impacts in different domains.

Even *within* a domain, we found that context matters. In their study of the features of YouTube, Lukoff et al. note that in most cases (77%) participants described video recommendations as reducing their sense of agency, but in some cases (23%) participants reported that recommendations actually supported their sense of agency by allowing them to lean back and let YouTube take control [73]. It depended on the the internal state of the user: whether they were visiting YouTube for a specific purpose or just to browse and pass the time. In short, damaging design patterns depend upon context and ACDPs are no exception. Therefore, our description of damaging patterns also capture the context in which they are most likely to lead to attentional harms.

Table 4 summarizes the typology of *11 attention capture damaging patterns* that we extracted from our literature review[11]. The rest of this section describes those 11 designs and reports a definition, context of use and examples for each.

3.2.1 **Infinite Scroll**. Infinite Scroll (N= 16, [6, 11, 13, 27, 29, 73, 75, 77, 82, 85, 91, 96, 104, 119, 128, 132]) is a design pattern in which, as the user scrolls down a mobile app or a website on their PC, more content automatically and continuously loads at the bottom. Despite its advantages, infinite scroll may become an "harmful feature" [85] or an "anti-pattern" [128] that promotes endless usage sessions [13, 128]. According to the studies in our corpus, the effects of patterns like Infinite Scroll can be *"understood or at least reasoned about in terms of established psychological theories"* (Notasil and Payne [91], p. 3). Infinite Scroll, in particular, can be related to the operant conditioning theory [113] and the variable reward technique [82] since it creates the illusion that new interesting content will "flow" forever. Unfortunately, the "quality" of the next visualized items cannot be predicted. Furthermore, Infinite Scroll is a good example of how attention capture damaging patterns automate interactions to reduce the individual's physical and mental effort to spend more time on the platform.

Context and Examples. Of the 16 papers we reviewed that mentioned Infinite Scroll, 14 were in the context of social media such as Facebook (9, i.e., [13, 27, 29, 75, 77, 85, 91, 104, 128]), Instagram (5, [29, 96, 104, 119, 128]), and Twitter (2, [11, 132]). Therefore, we

---

[11]The last 11 columns of Table 2 summarize the full list of papers that have been used to extract the typology.

Table 4: A typology of 11 attention capture damaging patterns.

Pattern Name	Description	Main Context(s) of Use
P1 - Infinite Scroll	As the user scrolls down a page, more content automatically and continuously loads at the bottom.	Social media (e.g., Facebook, Instagram, and Twitter).
P2 - Casino Pull-to-refresh	When the user swipes down on their smartphone, there is an animated reload of the page that may or may not reveal new appealing content.	Social media on smartphones.
P3 - Neverending Autoplay	A new video is automatically played when the current one finishes. There is never a point for the user to stop and reflect, and the option to turn off autoplay is hidden or non-existent.	Social media and video streaming platforms, e.g., YouTube.
P4 - Guilty Pleasure Recommendations	Personalized suggestions that prey on individual consumer frailty to target user's guilty pleasures and increase use time.	Social media and video streaming platforms, e.g., YouTube.
P5 - Disguised Ads and Recommendations	Advertisements and recommendations, e.g., posts and sponsored pages, that are disguised as normal content into social networks' newsfeeds.	Social media.
P6 - Recapture Notifications	Notifications that are deliberately sent to recapture users' attention and have them start a new usage session, e.g., notifications with recommended content or notifications about content the user has never interacted with.	Social media, video streaming platforms, and messaging applications.
P7 - Playing by Appointment	Users are forced to to use a digital service at specific times as defined by the service, otherwise the user may loose points and achievements.	Video games (mostly on social networks) and social media in general.
P8 - Grinding	Users are forced to repeat the same process several times to unlock an achievement, e.g., a new level in a video game or a badge on a social network.	Video games and social media.
P9 - Attentional Roach Motel	Registering to and accessing attention-capture digital services is easy, while operations like logout or canceling an account are painfully difficult.	Social media, e.g., Facebook.
P10 - Time Fog	A pattern through which designers reduce users' awareness of time spent, e.g., by hiding the smartphone's clock.	Video streaming platforms, e.g., Netflix.
P11 - Fake Social Notifications	The platform sends messages pretending to be another user or social notifications about some content the user has never interacted with.	Video games (mostly on social networks) and social media in general.

hypothesize that Infinite Scroll is most problematic in bite-size content, e.g., as in social networks. With Infinite Scroll, social networks provide unlimited new content to the user [73, 91], with the risk of making users *"passively slip into a dissociative state while scrolling"* (Baughan et al. [11]). Passively and mindlessly scrolling the newsfeed of a social network, in particular, negatively influence users' digital wellbeing [126], and it is one of the reasons why people feel nowadays conflicted about the amount of time they spend on their devices [66]. A participant of the study by Tran et al. [119], for example, reported that *"I go on Instagram and I just scroll through even though there's no real purpose."* Similarly, a participant of the study by Aranda et al. [6] said *"I hate when I spend time just scrolling and scrolling… it's all mind-numbing, and I don't benefit from any of it."*

3.2.2 **Casino Pull-to-refresh**. Casino Pull-to-refresh (N = 6, [20, 29, 34, 66, 91, 119]) is an interaction technique through which users can "pull" an interface, e.g., by swiping down on a mobile app, to reload the status of the system manually. As the user performs the swipe, there is an animated reload of the page, e.g., through a reload

wheel icon, that may or may not reveal new appealing content, e.g., an incoming email or a new friend's post. As the papers in our corpus and tech-insiders [68] warn, such a design pattern can be classified as an attention capture damaging pattern that offers a *variable reward* to its users. Indeed, it may result in a compulsive usage pattern that makes users repeatedly refresh an app hoping for new content to appear [6, 66]. In other words, pull-to-refresh exploits the same psychological vulnerabilities typically targeted in gambling addictions, e.g., in slot machines, since new rewards may be available at any time, e.g., messages or notifications.

Context and Examples. As reported by the 6 papers mentioning pull-to-refresh as an ACDPs, such an interaction technique characterizes touch-based interfaces, and targets social network users on smartphones (N = 5, [29, 34, 66, 91, 119]). According to Nontasil and Payne [91], animated pull-to-refresh techniques on social networks' mobile apps are deliberately modeled on "one-armed bandits." While we were expecting to find this pattern also in other contexts (see the email checking habits described by Oulasvirta et al. [94]) the fact that this pattern appeared predominantly on

social networks suggests that casino-like pull-to-refresh techniques are most problematic when the underlying content is various and less predictable than, for example, a simple email. In this way, not only the quantity (whether or not a reward is present), but also the quality of the reward is variable (the degree to which the new social media post(s) satisfy the user).

3.2.3 **Neverending Autoplay**. Neverending Autoplay is a design pattern in which new videos are continue playing indefinitely without any user interaction. Many papers included in our review (N = 14, [13, 16, 20, 26, 27, 29, 61, 73, 75, 77, 82, 96, 119, 128]) describe it as one of the most common attention capture damaging patterns. As with other patterns, autoplay can be a useful feature in some circumstances, e.g., to listen to YouTube's music videos while working, and detrimental in others, e.g., when it is used to attract attention against the user's best interests [73]. In particular, people's digital wellbeing problems arise when autoplay is "neverending" and cannot be easily turned off. Similarly to Infinite Scroll, indeed, Neverending Autoplay *"works by continuously providing users with yet another film clip for them to watch after one finishes â€" allowing 'endless' video streaming sessions"* (Widdicks et al. [128], p. 5). Bongard-Blanchy et al. [16] selected Neverending Autoplay as one of the deceptive designs to be investigated. While their recruited participants described autoplay as more acceptable than other strategies, e.g., hiding information, they also found it as one of the more influential features implemented by digital services to drive people's behavior. Lukoff et al. [73] classified autoplay as an ACDPs that undermines users' sense of agency, as it removes the need for autonomous decision-making [20].

Context and Examples. Nevereding Autoplay is an ACDP that is common in social networks (N = 5, [27, 29, 61, 77, 128]) and video streaming platforms (N = 6 , [13, 26, 29, 73, 75, 82]). As services with different characteristics, the pattern can be present in slightly different variations and with different goals. In social networks like Facebook (N = 3) and Instagram (2), videos embedded in the newsfeeds start automatically as long as they appear on the screen. At the same time, users' stories flow on their own or through a simple tap on the screen. Autoplay is also often active by default, and settings to deactivate it are often difficult to access [85], meaning that most users experience this pattern during all their usage sessions. YouTube (5), instead, attracts users by automatically (and infinitely) starting a new video [91] when the previous one ends (Figure 3). As reported by a participant of the study by Lukoff et al. [73], for instance, *"I often spend more time than I meant to because there is a good related video that seems worth watching so ya know, "Just one more' which becomes a couple hours."* It is worth noting that, unlike other services, YouTube users can easily disable/enable the autoplay functionality through a slider embedded in the video player. While Neverending Autoplay works differently depending on the underlying service, the common point is that there is never a stopping cue or pause for reflection for the user. In particular, videos on social networks are generally short and often consumed with little or no attention, e.g., while passively scrolling the newsfeed [126]. Thus, Neverending Autoplay on social networks is used to attract the user and maximize the amount of (different) content the user interacts with. On platforms like YouTube, instead, autoplayed videos

"follow" the previous ones as step-by-step recommendations, thus enforcing more extended viewing sessions [26].
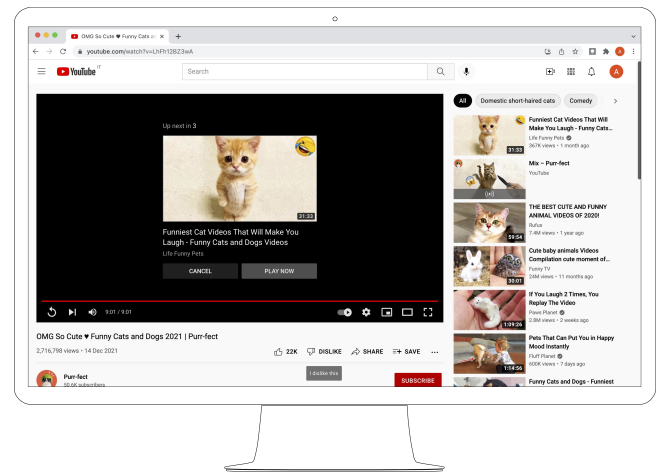

Figure 3: An example of the Neverending Autoplay pattern on YouTube. By default, a new video automatically starts after a 7-seconds countdown when the previous video ends.

3.2.4 **Guilty Pleasure Recommendations**. Guilty Pleasure Recommendations (N = 12, [6, 13, 20, 26, 29, 50, 73, 77, 91, 96, 114, 132]) are personalized suggestions that pray on individual consumer frailty to target every guilty pleasure of the users and keep them on the platform. They offer pleasurable content on some level but also leave users feeling guilty afterward. Many digital services use recommender systems to propose new and appealing content to the user based on their past interactions (content-based approach) or the preferences of similar users (collaborative filtering approach). Recommendations are undoubtedly an important mechanism that can improve the overall user experience with a platform that is designed to maximize a user's utility [20]. However, as reported by the 12 papers under analysis, misalignment between the goals of the platform and the user's goals – i.e., a value alignment problem [20] – can make recommendations an attention capture damaging pattern that "trap" the users in the system and keeps their attention [13, 114]. These clickbait suggestions [73, 77] increase the platform's utility without a benefit for the user. In particular, the paper by Chaudhary et al. [26] talks about "bias grind," by referring to UI patterns that *"disproportionately overload user interests and biases [...] providing an infinitely long scroll of Recommendations based on previous watching history"* (p. 788). Unfortunately, Guilty Pleasure Recommendations cannot be easily personalized or disabled without third-party tools, e.g., Unhook [123]. As for other ACDPs, *"the variable schedule of rewards in content recommendations also play a huge role in hooking users"* (p. 456:20), as reported by Cho et al. [29]. Furthermore, Guilty Pleasure Recommendations are particularly harmful to people lacking self-control and self-esteem [13, 50].

Context and Examples. The 12 papers mentioning Guilty Pleasure Recommendations as an ACDP are in the context of social media, e.g., Facebook (4, [13, 29, 77, 91]), and video streaming platforms, e.g., YouTube (3, [13, 29, 73]). Both social media and video streaming

platforms often provide their users with clikbait [73, 91] suggestions, thus increasing compulsiveness in their long-term usage [26]. Video streaming platforms like YouTube display an unlimited number of personalized (often viral) suggestions in almost every part of their interfaces, with the main aim of attracting users' attention and making them watch more videos [13] (Figure 4). According to the study by Chaudhary et al. [26], recommendations in these platforms may be used to extend users' current viewing sessions and increase the chances of experiencing regret by 34%. In contrast, social networks can recommend different kinds of content, from friends to follow [13] to games [77] and trending topics [132]. In some cases, these recommendations are also deliberately disguised into the users' newsfeed (see the Disguised Ads and Recommendations pattern). Furthermore, the study by Cho et al. [29] highlighted that social networks like Instagram place recommendation-based features close to features that are used more actively, e.g., the search bar, thus causing *"habitual feature tour and sidetracking from the original intention of app use"* (p. 1). All in all, the papers under analysis suggest that Guilty Pleasure Recommendations, both on social networks and on video streaming platforms, are particularly harmful when they are frequently updated to increase the platform's utility [20]. Lukoff et al. [73] cited a study by Pew Research that finds that YouTube's recommender system directs users towards progressively longer videos [110]. In discussing the study by Bakshy et al. [9], instead, Burr et al. [20] noted that Facebook uses the users' feedback to refine its future recommendations and provide users with more catchy suggestions.
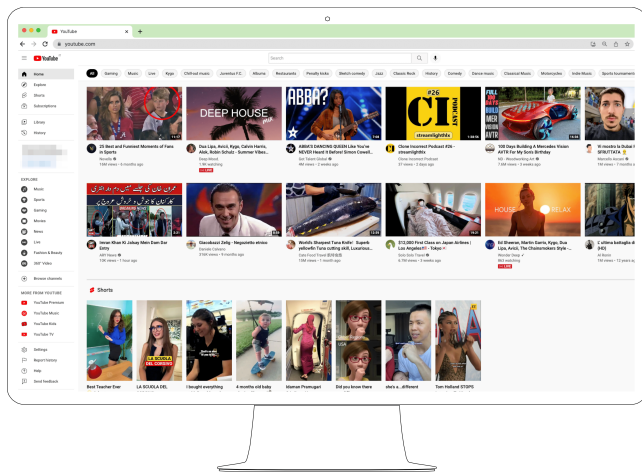

Figure 4: An example of the adoption of Guilty Pleasure Recommendations on YouTube. The figure shows the typical homepage of the video streaming platform: it provides users with several personalized (viral) content that primarily aims to attract them to click and watch "just another video."

3.2.5 ***Disguised Ads and Recommendations****. The *www. deceptive.design* website defines the Disguised Ads pattern as *"adverts that are disguised as other kinds of content or navigation, in order to get you to click on them."* This deceptive design has already been investigated by previous works exploring dark patterns

in websites (e.g., [47]) and mobile apps (e.g., [33]) and has been classified as a form of interface interference [46]. In our work, we extend the original definition of the Disguised Ads pattern by referring to the practice of mixing disguised advertisements with recommendations that are camouflaged as normal content (N = 14 , [20, 29, 33, 39, 41, 44, 73, 77, 85, 98, 114, 125, 131, 132]).

Context and Examples. Disguised Ads and Recommendations is an ACDPs that is typically used to increase the time users spend on social networks. Indeed, all the 14 papers in our corpus that mentioned this pattern were in the context of social media, e.g., Facebook (6, [20, 73, 77, 85, 98, 114]) and Twitter (3, [11, 98, 132]). These digital services purposely inject new, catchy content resembling friends' posts in their newsfeeds to mislead users to click it more often [44]. Injection may not only involves advertisements [41, 44, 131], but also sponsored pages [33, 98] and recommended posts [29, 77], often by other people the user does not follow. For example, sponsored pages on Facebook and Instagram are mixed with stories and posts from friends and followed pages (Figure 5): when clicking on them, the user is still using (and paying attention to) the same service. In this sense, this kind of injection can be seen as a particular type of personalized recommendation – delivered according to the specific user's profile – that influence people's behavior online [114]. Another distinctive example of the Disguised Ads and Recommendations pattern can be found on Twitter, that often displays tweets from people that the user is not following, e.g., tweets of users followed by a friend or generic *"you might like"* tweets, by pretending they are normal content. Mildner et al. [85] highlight that "smart" social media newsfeeds with Disguised Ads and Recommendations are tempting and likely increase the chances of prolonging usage sessions, thus causing a sense of regret. As warned by Burr et al. [20], the pattern of Disguising Ads and Recommendations into social networks is also problematic because newsfeeds become *"a representation of what the ISA expects will elicit the most clicks based on prior behaviour"* (p. 755), rather than a representation of the users' belief and preferences. Unfortunately, most users are not able to process such a misalignment. Susser et al. [114], for example, referenced a survey by Pew Research highlighting that more than half of Facebook users do not understand well why certain posts are included in their newsfeed and others are not.

3.2.6 ***Recapture Notifications****. Recapture Notifications (N = 7, [6, 11, 13, 73, 77, 125, 132]) are notifications that are deliberately sent to recapture the attention of a user who escaped or left a digital service for some period of time, e.g., to make the start a new usage session. Several previous works have studied the influence of notifications on users' digital wellbeing. Indeed, the huge and continuously growing number of notifications that users receive every day [97] can interfere with daily activities like studying, working, and driving [4, 48], and make users less productive [79] and more stressed [80]. As a consequence, several digital self-control tools nowadays help users filter or block notifications [87]. While notifications can alert users to important information, the analyzed papers highlight that Recapture Notifications are an attention capture damaging pattern that should be managed or avoided, as they used as a pretext to make user unlock a device and going into apps or websites to engage further [6]. According to a participant in
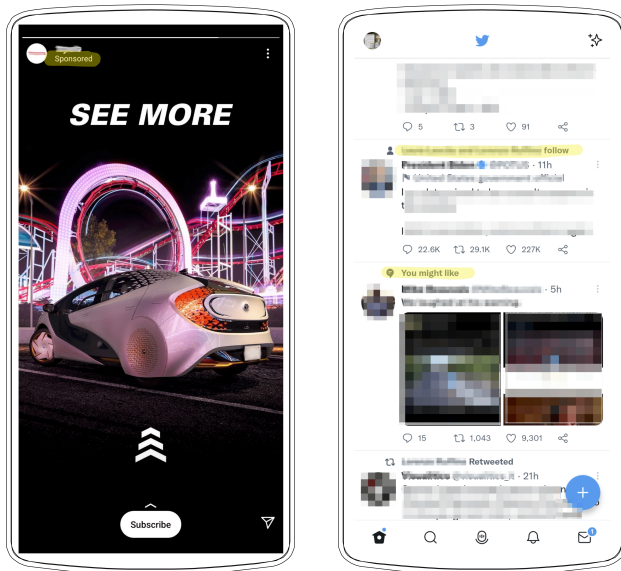
Figure 5: Two examples of the Disguised Ads and Recommendations pattern. In (a), Instagram's sponsored content resembles a story from a friend, and is inserted between other stories. The sponsored badge (on the top-left corner) is tiny and barely visible. In (b), there are two tweets from people that the user is not following: the first one is a tweet from a "friend of a friend," the second one is displayed under a generic "you might light" badge.

Lukoff et al. [73], for example, Recapture Notifications *"draw me to YouTube and create my schedule for 20-30 minutes, this creates an addiction"* (p. 7). Unfortunately, these notifications are typically activated by default in contemporary digital services [10] and often distract users [6, 13, 61].

Context and Examples. In the papers we analyzed, sending Recapture Notifications is a cross-cutting design pattern that characterizes social media (4, [11, 13, 77, 132]), video streaming platforms (2, [13, 73]), and instant messaging applications (1, [125]). Specifically, the content of notifications plays an important role in determining their (negative) impact [73], and Recapture Notifications typically convey unimportant unimportant information [6]. A classic example of Recapture Notifications are ones that share information about others' activities on social networks [77]. A participant in the study by Lyngs et al. [77], for instance, stated that *"if I didn't have things popping up every 30 minutes like 'this has happened' I don't think I would think about Facebook"* (p. 8).

3.2.7 *Playing by Appointment*.
Playing by Appointment (N = 5, [41, 73, 82, 128, 130]) is an attention capture damaging pattern that forces users to use a digital service at specific times as defined by the service, rather than the user [130]. It has been classified as a "temporal dark pattern" by Zhagal et al. [130], that related it to a time-wasting activity that attempts to test the user's patient [41]. The pattern is engineered to encourage users to re-visit a digital service to avoid losing the possibility of earning something, e.g., points or even the ability to progress in a game.

Context and Examples. Playing by Appointment has been originally studied in the context of social media games [130], wherein resources may whither away if the user does not access the game at specific times [73]. Zagal et al. [130], for example, mention the game FarmVille. In this social media game, users that plant crops are encouraged to return to the game after a given amount of time not only because they can earn points but because a crop that is not harvested in time loses its value. Zagal et al. [130] also mention "lighter" versions of Playing by Appointment, e.g., as in some Pokémon games. Here, some Pokémons can only be captured at specific hours of the day, but a player can complete the game even without capturing them. Besides games, we also highlight the possibility of finding the Playing by Appointment pattern in social networks. An example is the Snapchats' social streaks, which count how many consecutive days two people have been sending Snaps to each other [28]: keeping up a Snapchat streak gives the user extra points, while even a single day without sending a Snap breaks the streak. Another example of Playing by Appointment can be found on the BeReal social network [14], which asks users to publish every day a post at a time that is randomly selected by the system and communicated to the user through a notification.

3.2.8 *Grinding*.
Grinding (N = 7, [33, 41, 46, 47, 82, 128, 130]) is an attention capture damaging pattern that forces users to repeat the same process several times to unlock an achievement [33]. As Playing by Appointment, grinding has been classified as a "temporal dark pattern [130]". According to Widdicks et al. [128] in the context of video games, *"it is interest of a game developer to entice players to commit more time to a game than what the player expects or plans, encouraging players to waste time"* (p. 2). Through Grinding, digital services "consume" the user's time and attention by increasing engagement and promising a later achievement [47], e.g., a new level in a video game or a badge on a social network. Di Geronimo et al. [33], reports that identifying this kind of damaging patterns is not easy, as they are initially disguised as features that increase user engagement.

Context and Examples. Grinding is an ACDPs that has been defined by Zagal et al. [130] in the context of their research work on "dark game design patterns". According to the authors, Grinding is common in massively multiplayer games, e.g., World of Warcraft. Here, players are forced into *"needlessly spending time in a game for the sole purpose of extending the game's duration"* (Zagal et al. [130], p. 3), e.g., killing monsters to gain experience points [46]. Besides multiplayer games, researchers highlight that Grinding can be adopted in social media games [46], e.g., FarmVille, and social networks in general. An example is the verified badge on Twitter, which can be achieved if the account is *notable* and *active*, e.g., with a sufficient number of followers and mentions [122].

3.2.9 *Attentional Roach Motel*.
Attentional Roach Motel (N = 6, [10, 13, 33, 47, 78, 85]) represents the (engineered) difficulty of canceling an account or logging out from an attention-capture digital service, in contrast to the simplicity of creating an account and accessing the service. This deceptive design pattern is an extension of the original Roach Motel pattern described by the *www.deceptive.*

design website and previous typologies, e.g., [33, 46, 47]. It was defined as a mechanism that generates situations for the user that are easy to get in, but hard to get out. Besides entrapping users into paid subscriptions, Baroni et al. [10] highlighted that tech companies can also use the Roach Motel dark pattern to keep users' attention by keeping them as customers of their services, e.g., by depriving users of the possibility of deleting their accounts [10, 33, 78, 107]. Similarly, the Attentional Roach Motel pattern may be exploited to make account settings difficult to access [85], e.g., by relegating them to small drop-down menus, thus hindering the possibility of logging out from a digital service [33, 47, 78, 85]. As discussed by Mildner and Savino [85] in their study of Facebook's damaging design patterns, moving logout buttons into drop-down menus can be considered as an interface interference [46]. All in all, UIs adopting the Attentional Roach Motel pattern affect how alternatives are perceived by promoting a predefined action. A way to hide available settings, in particular, is to use deceptive visualizations that leverage the *salience bias* [121], e.g., to create optical illusions and alter people's perceptions of the different buttons on the user interface.

Context and Examples. Overall, 3 out of the 6 papers under analysis, i.e., [13, 78, 85], found the Attentional Roach Motel pattern on Facebook, thus suggesting that this pattern is most problematic and evident on social media. By exploring the effects of damaging patterns on Facebook, for example, Mildner and Savino [85] found that the logout button was moved from the top-navigation bar into the 'Account' menu in 2010, thus limiting discoverability. Bhoot et al. [78], instead, described the overwhelming process of deactivating or deleting a Facebook account: 1) searching for the settings, 2) finding the 'Deactivation' tab (included in the 'Your Facebook Information' tab), 3) choosing between 'Deactivating' and 'Permanently deleting' the account, 4) entering the user's password, 5) inserting a reason for deactivation/cancellation, and 6) skipping a final pop-up dialog suggesting to 'log out' instead of deactivating/canceling the account. As reported on the Facebook website [38], moreover, users' information is canceled from the platform after 30 days, only. Furthermore, a login during the 30 days following deletion allows the user to re-activate the account. In the case of a deactivated account, a login anytime after the deactivation will automatically reactive the account.

3.2.10 **Time Fog**. Time Fog (N = 1, [26]) is an attention capture damaging pattern through which designers deliberately *"induce unawareness by reducing autonomy of monitoring user time spent"* (Chaudhary et al. [26], p. 785). With respect to the original name given by Chaudhary et al. in their investigation of damaging patterns on video streaming platforms, i.e., "feature fog," we decided to use the word "time" to indicate further that the pattern is specifically about obscuring users' sense of time spent. The goal of this pattern is to reduce the possibilities for users to get feedback on the time they spend on digital services, e.g., by hiding the video elapsed time, thus increasing the chances of longer usage sessions. According to Chaudhary et al. [26], Time Fog is related to the "hidden information" pattern reported in the *www.deceptive.design* website, as well as the "interface interference" described in the Gray et al. taxonomy [46]. Therefore, as for the Attentional Roach Motel, Time Fog can be considered a deceptive visualization that leverages the

salience bias [121]. Chaudhary et al. also highlighted a similarity with the 'menu engineering' trick [65], through which restaurants hide costly items in the menus so that they are not directly visible to customers.

Context and Examples. The only paper mentioning Time Fog as an ACDP was in the context of video streaming platforms. One of the examples reported by Chaudhary et al. [26], in particular, was about Netflix. The authors noted that *"the time elapsed feature that lets a user monitor how much time has elapsed since the start of video is missing from Netflix"* (p. 785). Such a feature enforces extended viewing sessions because users cannot easily tell how much time is left until the end of a video. Besides video streaming platforms, another possible example of Time Fog can be found on mobile games that typically start full-screen by hiding the smartphone's clock.

3.2.11 **Fake Social Notifications**. Fake Social Notifications (N = 9, [10, 41, 46, 47, 61, 78, 125, 130, 132]) refers to the practice of deceiving users with false social activities and information. Following this ACDP, digital services may send messages on behalf of a user [41, 46, 130], e.g., by pretending to be them. According to Fitton et al. [41], these messages can be related to the Brignul's "Friends Spam" deceptive pattern. Similarly, digital services may communicate to a user a social activity of another person about content the user has never interacted with [61, 132]. All in all, these deceptive techniques violate the expectation that the received messages should actually be from a real person, and are often designed to spur the user receiving the message to open (and start using) a given digital service. Furthermore, being related to social activities, Fake Social Notifications may leverage on our *herd instinct bias* of replicating others' actions [30] as well as on the *spotlight effect* [43, 84], i.e., an egocentric bias that lead us to perform behaviors that elicit social approval.

Context and Examples. According to the 9 papers under analysis, Fake Social Notifications are common in games hosted on social networks (4, [41, 46, 82, 130]) and on social networks themselves (2, [61, 132]). Games like Farmville and the Candy Crush Saga, for example, may *"impersonate other players by communicating actions they never performed, thus misleading the player about the activities of their friends in the game"* (Zagal et al. [130], p. 6). Games on social networks may also send invitations to join the community to the player's friends [41] by spamming all the players' contacts through messages that claim to be from the player [46]. Regarding social networks, Kollnig et al. [61] and Zhang et al. [132] highlighted the presence of Fake Social Notifications on Twitter, e.g., when the platform sends the notification *"user x just tweeted after a while"* (Figure 6a). Finally, an instance of this pattern may also be found in some instant messaging applications, e.g., Telegram, that broadcast messages like *"user x just joined, say hello"* (Figure 6b).

4 DISCUSSION

4.1 Damaging Design Patterns: A New Name for Dark Patterns

Inclusive language matters in creating a welcoming culture in the field of human-computer interaction, which still excludes many

Figure 6: Two examples of Fake Social Notifications. In (a), Twitter deliberately notify the user that a contact has just tweeted something after a while, assuming that the content would be of interest to the user. In (b), Telegram incentivizes the user to send a message to a new contact without asking either party for confirmation.

from participation [93]. In recent years, many organizations have moved on from the oppressive terminology endemic to our field, for example from "master/slave" to "parent/child" and from "blacklist" to "block list" [124]. In the case of "dark patterns," the association of "dark" with harm is problematic as it may reinforce the racist heuristic of viewing people with darker skin tones as evil, also known as the "bad is black" effect [3]. It is time for a new name.

Our initial response was to adopt the term "deceptive design patterns," following the lead of Harry Brignull (who renamed his website from *darkpatterns.org* to *deceptive.design*) and others who have also made this switch [55, 108, 129]. And "deceptive" largely works for Brignull's original 12 patterns, which mostly operate by tricking the user into a false belief. However, we came to realize that most of the dark patterns that design practitioners and researchers have identified since then are not actually deceptive per se. Take for example the Pay to Skip pattern in [130], in which players have to pay more to bypass a tedious or difficult part of a game – there's no real deception involved. Only 22 of 86 (25%) dark patterns were coded with deception as a required or optional attribute in a recent review by Mathur et al. They write, *"It is difficult to see where the trick is, however, in Brignull's Confirmshaming or Bösch et al.'s Forced Registration – these user interface designs are often entirely (and frustratingly) transparent to users"* [82] (p.7). Bongard-Blanchy et al. note that even though users are generally awareof dark patterns, they are still unequipped to resist their influence, which calls into question that idea that deception is the primary mechanism by which dark patterns operate [16]. Similarly, in our own literature

review, we quickly discovered that some attention capture patterns operate by deception, but not all.

We therefore advance the term "damaging pattern" as an inclusive alternative to "dark patterns," as proposed in a blog post by the researcher and artist Caroline Sinders, who herself cites an unnamed workshop participant [108]. Some view the 'vagueness' of the term dark patterns as a weakness, but it can also be seen as a strength. Its 'broadness' has oriented designers, researchers, and regulators towards harms created by interface design – even if much work remains to identify specific patterns and specific harms. Like dark patterns then, the term "damaging pattern" suggests an openness to a wide variety of design patterns that intentionally diminish the user's wellbeing.

4.2 Deception and Seduction

In our own typology, attention capture damaging patterns fell into two categories: *deceptive designs* (4/11) and *seductive designs* (7/11). Designs that used deception were Disguised Ads and Recommendations, Attentional Roach Motel, Time Fog, and Fake Social Notifications. This set of patterns are often strong candidates for regulation, perhaps because it is feasible (although not easy) to set standards for what constitutes an interface that tricks the user into a false belief. For example, when it spammed users with invite emails from 'friends' that those friends never sent, LinkedIn was forced to pay a settlement of $13 million for its Fake Social Notifications [90]. Similarly, social media platforms are required to disclose advertising content, although they do often find ways to make the distinction between paid and organic as subtle as possible (i.e., Disguised Ads).

However, the other seven ACDPs we identified might be called *seductive designs*, in that they tempt the user with short-term satisfaction. For example, Guilty Pleasure Recommendations leverage recommender systems to deliver temptations that are designed to exploit psychological vulnerabilities. Being seductive, the "effectiveness" of these patterns may also depend on the user that is experiencing them, e.g., their current mood or level of self-control. As with attention capture itself, seductive designs is a category that we found to be largely neglected in the current scholarship. In the Mathur et al.'s review [82], seduction is included in only one of 19 definitions and is not included at all in their higher-level attributes of damaging designs.

This lack of recognition for seductive designs may be due to concerns that it could take away from more unambiguously harmful patterns. Of course, 'drawing the line' between the acceptable, the tolerable, and the truly damaging is universal challenge for design patterns [46], but it does seem particularly challenging in the case of seduction. Users rarely if ever desire to be deceived, but they do sometimes wish to be seduced. There is perhaps some overlap with health policy: regulations against lead in water are widely accepted by the public, but proposals to impose soda taxes are highly controversial [92]. Similarly, users have different notions of what and when experiences count as guilty pleasures and just how much is too much [72]. Faced with this ethical quagmire and the business incentive to maximize user time on site, it is easy for designers to set the default to 'unlimited guilty pleasure.'

So how *can* designers determine how much seduction is too much? Luguri and Strahilevitz propose A/B testing as a general

approach for evaluating damaging designs [70], e.g., comparing the sign up rate for a dubious subscription service with and without a damaging design pattern. However, in the case of attention capture, the metrics are less clear: an increase in time spent on an app does not necessarily imply harm to the user [74]. Instead, quantitative measures like 'screen time' might be triangulated with other qualitative measures of digital wellbeing [71], such as the three impacts included in our definition of ACDPs: lack of goal awareness, lack of sense of time and control, and a sense of regret. In fact, Cho et al. recently innovated a promising technical approach for tying the use of specific features within apps (e.g., the Instagram newsfeed) to qualitative measures (e.g., regret as measured by experience sampling) [29]. Approaches like this may help designers identify when a seductive design goes too far.

4.3 ACDPs Across Domains and Interfaces

The systematic literature review described in this paper revealed a typology predominantly related to a specific domain, i.e., social media and gaming (mostly games on social networks). This narrow focus is not surprising, as social networks have been described as a potential threat to users' digital wellbeing by several previous works, e.g., [109, 125, 126]. However, we believe that some (if not all) of the presented patterns are already adopted or could be potentially applied across different domains. As described in Section 3.2, Recapture Notifications is a pattern that characterizes video streaming platforms and messaging applications in addition to social networks. We can also find the same pattern in other domains, e.g., educational technologies. Duolingo [36], for example, is a language learning tool that makes extensive use of push notifications to remind users to practice a foreign language. Other patterns not included in our typology, e.g., the "Limited-time Message" and "Countdown Timer" patterns described by Mathur et al. [81] in the context of Shopping websites, may become an ACDP in other domains, e.g., by applying time urgency to social messages from friends that "expire" or disappear. We also believe and see the opportunity to translate the concept of ACDP into the work/productivity domain, a context in which the attentional harms produced by digital services is often neglected. As reported by Cecchinato and Cox in their critical review on boundary management and communication technologies [24], current technology has the potential to undermine people's work-life boundaries by requiring their constant attention. Indeed, while disconnection is fundamental to recovering from work-related stress, devices like smartphones often force people to be constantly connected in an attempt to increase productivity and satisfy employers. Future works could try to map our typology to work-related technologies and tools and see which patterns are most common and harmful in this domain.

Another aspect that emerges from our work is that researchers have traditionally associated damaging patterns with graphical user interfaces. Although most digital services are typically accessed through a "visual" interface, we highlight that attention capture damaging patterns may work and be adopted across various existing and future interfaces. For example, Amazon Alexa can autonomously deliver "by the way" vocal suggestions to recommend users adopt specific commands. These suggestions can be seen both as recommendations disguised as normal messages and as a Recapture Notifications.

Overall, we do not see our typology as a fixed and final set of design patterns but rather as a starting point to encourage researchers and practitioners to investigate the concept of ACDPs in other domains and for various interfaces.

4.4 Implications for Designers

Our work has the potential to inform new design processes that prioritize users' digital wellbeing as a top design goal. By leveraging the definitions and examples of attention capture damaging patterns reported in this paper, the HCI community and tech providers could work together to find alternative business models and design processes that do not necessarily target users' attention and engagement: designing for users' digital wellbeing, e.g., by targeting meaningful interactions [75] and microplanning [73], may initially result in a lower user engagement, thus resulting in lower business profitability in the short term, but it could increase user loyalty in the long term.

4.4.1 Evaluating existing digital interfaces for attentional harms. Among other potential uses, we envision our typology as a tool to help evaluate existing products and digital interfaces and adopt proper countermeasures. The ACDPs in our typology and related examples could be used as a check list for designers to audit their own designs for attentional harms. In parallel, designers could conduct A/B experiments to detect the presence of possible attention capture damaging patterns in their interfaces [70] and mitigate their negative influence on users' digital wellbeing. Some important tech companies have recently adopted similar experiments. For example, Instagram recently tested removing the "like" count in several countries. The CEO of Instagram explained that the aim of these experiments was to "*depressurize*" Instagram for young people by "*making the platform less of a competition and give people more space to focus on connecting with the people that they love* [111]." Our typology could help designers and design researchers identify potentially damaging features to evaluate in such experiments.

4.4.2 Promoting the Adoption of Alternative Designs That Respect User Attention. Beyond identifying attentional harms, our work could help designers by directing them towards approaches and patterns that respect users' time and attention. Some of these are conceptual design approaches raised in the prior literature we reviewed. For example, in their analysis of the interaction between people and intelligent software agents, for example, Burr et al. [20] suggest computing recommendations based on the user's past problematic behaviors, e.g., to avoid catching suggestions for users that already showed signs of addictive behaviors. Similarly, Lukoff et al. [73] highlighted the need to rethink the concept of "relevance" for recommender systems by considering when recommendations might promote excessive and compulsive usage of the service.

We also highlight several concrete examples of countermeasures and alternative design patterns for some the ACDPs that we identified in this work:

- Offer options for users to customize or disable features that they find distracting. For example, in YouTube, users can easily disable/enable the autoplay functionality. This is particularly applicable to features whose effects vary by situation, as is the case with Neverending Autoplay [73].
- Highlight the estimated time investment of new content, so that users can avoid opening an app or a website if it does not fit their current situation. For example, the top of each blog post on the site Medium[12] displays an estimated read time. This design contrasts with the Time Fog pattern, which purposefully obscures any indicator of the passage of time.
- In cases where there is content overload such as Infinite Scroll, enable the user to save content for future consumption and reduce the urge to consume everything now. For example, the Watch Later playlist that is built into YouTube lets users control when and where to watch videos, e.g., by saving videos on the go and watching them later at home.
- Ensure that advertisements are relevant, transparent, and clearly distinguishable from other content so they can be easily ignored. The AcceptableAds[13] framework, for example, delivers respectful, non-intrusive, and relevant advertisements that are compliant with criteria defined by an independent committee, e.g., ads must be marked with the word "advertisement" or its equivalent. Comparable criteria might also be applied to the related ADCP we identified: Disguised Recommendations, in which system recommendations are displayed in a way that is misleadingly similar to organic content such as posts from friends.

Overall, we recognize that an outright ban on the ADCPs we identified across all situations would be problematic. Indeed, some of the patterns we describe can be useful in certain circumstances such as when creating an immersive "flow" experience, e.g., when YouTube's autoplay is used to listen to music videos in the background. We therefore endorse the vision that emerges from the work of Lukoff et al. on YouTube internal mechanism [73], according to which "*the message of attention capture dark patterns should not be 'never X,' but rather 'be careful when X'*" (p. 14).

4.4.3 Developing tools for digital self-control. Besides evaluating and designing digital interfaces, the typology may also help inspire the design and development of novel and more efficient digital self-control tools that take a step beyond simple usage timers and lockers. By knowing how attention-capture damaging patterns operate, e.g., which users' psychological vulnerabilities they exploit, designers of these tools might directly target these patterns by developing theoretically grounded mitigation strategies. For instance, designers and researchers could explore the adoption of ACDPs nudges. Nudges can be defined as changes in the design architecture of a system that target users' cognitive biases [116]. Used in different domains, one of their main goals is to allow users to know the underlying system better and make deliberate decisions, i.e., they can be seen as a way to make users exercise their own agency [51]. In our context, nudges could be used to promote awareness of ACDPs and their negative consequences on people's digital wellbeing. Examples may include informative splash screens listing all

---

[12]https://medium.com/, last visited on November 19, 2022
[13]https://acceptableads.com/, last visited on November 19, 2022

the ACDPs exploited by a given mobile app, or widgets highlighting when an ACDP is operating, e.g., during intensive scrolling. Overall, grounding the design of behavior change technologies on well-established behavioral theories is fundamental to generate long-lasting results [99]. Nevertheless, such a theoretical-grounded approach still needs to be further established. Contamination between the HCI and behavioral science communities has already been successful in several fields, from supporting healthy diets [101] to promoting physical activities [15]. However, researchers highlighted a general lack of theory in digital wellbeing and digital self-control tools [76].

4.5 Limitations and Future Work

Our research has several limitations and also suggests opportunities for future work. First, digital wellbeing and damaging patterns are recent and evolving research areas, and new work is constantly emerging. Consequently, the corpus of papers included in our systematic literature review may soon become outdated. Second, we made subjective judgments as to which inclusion and exclusion criteria to use in selecting papers (e.g., in choosing relevant keywords). However, by transparently sharing our assumptions and methods, we hope that other researchers can easily evaluate and replicate our work. Third, our review focused primarily on the Computer Science literature. In particular, the majority of the papers included in our corpus are from human-computer interaction venues and journals. Researchers in other areas, from law to design and social sciences, could refine and extend our typology with additional studies and reviews.

5 CONCLUSIONS

As tech companies design their digital services to maximize time spent and daily visits, there is a need to better understand the design patterns that "steal" attention and harm digital wellbeing. To pursue this goal, we presented the results of a systematic literature review that provided a comprehensive overview of attention capture damaging patterns (ACDPs). Our typology of 11 ADCPs complements existing collections of damaging patterns that have been proposed in the domains of finance and privacy and is a call for designers and regulators to seriously consider attentional harms and the design patterns that promote them. We hope that the definition, criteria, and typology of ADCPs in this work serves as a common reference for technologists and policymakers aiming to align technology design with digital wellbeing.

REFERENCES

[1] 90 Percent Of Everything 2014. Some Dark Patterns now illegal. https://www.90percentofeverything.com/2014/08/26/some-dark-patterns-now-illegal-in-uk-interview-with-heather-burns/ Accessed: 2019-4-8.
[2] Christopher Alexander, Sara Ishikawa, and Murray Silverstein. 1977. *A Pattern Language: Towns, Buildings, Construction.* Oxford University Press, New York.
[3] Adam L. Alter, Chadly Stern, Yael Granot, and Emily Balcetis. 2016. The "Bad Is Black" Effect: Why People Believe Evildoers Have Darker Skin Than Do-Gooders. *Personality and Social Psychology Bulletin* 42, 12 (2016), 1653–1665. https://doi.org/10.1177/0146167216669123
[4] Morgan G. Ames. 2013. Managing Mobile Multitasking: The Culture of IPhones on Stanford Campus. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA) *(CSCW '13).* Association for Computing Machinery, New York, NY, USA, 1487–1498. https://doi.org/10.1145/2441776.2441945

[5] Hilary Andersson. 2018. Social media apps are 'deliberately' addictive to users. https://www.bbc.com/news/technology-44640959 Accessed: 2021-12-27.

[6] Julie H. Aranda and Safia Baig. 2018. Toward "JOMO": The Joy of Missing out and the Freedom of Disconnecting. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Barcelona, Spain) *(MobileHCI '18)*. Association for Computing Machinery, New York, NY, USA, Article 19, 8 pages. https://doi.org/10.1145/3229434.3229468 ✓.

[7] Dan Ariely. 2008. *Predictably irrational: The hidden forces that shape our decisions.* HarperCollins Publishers.

[8] American Psychological Association. 2021. APA Dictionary of Psychology. https://dictionary.apa.org/attention Accessed: 2021-12-27.

[9] Eytan Bakshy, Solomon Messing, and Lada A. Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132. https://doi.org/10.1126/science.aaa1160

[10] Luiz Adolpho Baroni, Alisson Andrey Puska, Luciana Cardoso de Castro Salgado, and Roberto Pereira. 2021. Dark Patterns: Towards a Socio-Technical Approach. In *Proceedings of the XX Brazilian Symposium on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, Article 15, 7 pages. ✓.

[11] Amanda Baughan, Mingrui Ray Zhang, Raveena Rao, Kai Lukoff, Anastasia Schaadhardt, Lisa D. Butler, and Alexis Hiniker. 2022. "I Don't Even Remember What I Read": How Design Influences Dissociation on Social Media. In *CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 18, 13 pages. https://doi.org/10.1145/3491102.3501899 ✓.

[12] Eric P. S. Baumer, Rui Sun, and Peter Schaedler. 2018. Departing and Returning: Sense of Agency as an Organizing Concept for Understanding Social Media Non/Use Transitions. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 23 (Nov. 2018), 19 pages. https://doi.org/10.1145/3274292

[13] Mohammed Bedjaoui, Nadia Elouali, and Sidi Mohamed Benslimane. 2018. User Time Spent Between Persuasiveness and Usability of Social Networking Mobile Applications: A Case Study of Facebook and YouTube. In *Proceedings of the 16th International Conference on Advances in Mobile Computing and Multimedia* (Yogyakarta, Indonesia) *(MoMM2018)*. Association for Computing Machinery, New York, NY, USA, 15–24. https://doi.org/10.1145/3282353.3282362 ✓.

[14] BeReal. 2022. BeReal. Your Friends for Real. https://bere.al/en Accessed: 2022-09-08.

[15] Agon Bexheti, Anton Fedosov, Jesper Findahl, Marc Langheinrich, and Evangelos Niforatos. 2015. Re-Live the Moment: Visualizing Run Experiences to Motivate Future Exercises. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (Copenhagen, Denmark) *(MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 986–993. https://doi.org/10.1145/2786567.2794316

[16] Kerstin Bongard-Blanchy, Arianna Rossi, Salvador Rivas, Sophie Doublet, Vincent Koenig, and Gabriele Lenzini. 2021. "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* (Virtual Event, USA) *(DIS '21)*. Association for Computing Machinery, New York, NY, USA, 763–776. https://doi.org/10.1145/3461778.3462086 ✓.

[17] Christoph Bösch, Benjamin Erb, Frank Kargl, Henning Kopp, and Stefan Pfattheicher. 2016. Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns. *Proceedings on Privacy Enhancing Technologies* 2016, 4 (2016), 237–254. https://doi.org/10.1515/popets-2016-0038

[18] Harry Brignull. 2016. Dark Patterns. https://www.youtube.com/watch?v=zaubGV2OG5U

[19] Harry Brignull. 2020. What is deceptive design? https://www.deceptive.design/ Accessed: 2021-12-27.

[20] Christopher Burr, Nello Cristianini, and James Ladyman. 2018. An Analysis of the Interaction Between Intelligent Software Agents and Human Users. *Minds Mach.* 28, 4 (dec 2018), 735–774. https://doi.org/10.1007/s11023-018-9479-0 ✓.

[21] Christopher Burr, Mariarosaria Taddeo, and Luciano Floridi. 2020. The Ethics of Digital Well-Being: A Thematic Review. *Science and Engineering Ethics* 26 (2020), 2313–2343. https://doi.org/10.1007/s11948-020-00175-8

[22] Lisa Butler. 2006. Normative Dissociation. *The Psychiatric clinics of North America* 29 (04 2006), 45–62, viii. https://doi.org/10.1016/j.psc.2005.10.004

[23] Ana Caraban, Evangelos Karapanos, Daniel Gonçalves, and Pedro Campos. 2019. 23 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300733

[24] Marta E. Cecchinato and Anna L. Cox. 2020. Boundary Management and Communication Technologies. In *The Oxford Handbook of Digital Technology and Society.* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190932596.013.10

[25] Rory Cellan-Jones. 2018. Confessions of a smartphone addict. https://www.bbc.com/news/technology-44972913 Accessed: 2021-08-06.

[26] Akash Chaudhary, Jaivrat Saroha, Kyzyl Monteiro, Angus G. Forbes, and Aman Parnami. 2022. "Are You Still Watching?": Exploring Unintended User Behaviors and Dark Patterns on Video Streaming Platforms. In *Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 776–791. https://doi.org/10.1145/3532106.3533562 ✓.

[27] Justin Cheng, Moira Burke, and Elena Goetz Davis. 2019. Understanding Perceptions of Problematic Facebook Use: When People Experience Negative Life Impact and a Lack of Control. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 1–13. ✓.

[28] Childnet. 2017. What is a Snapchat streak? https://www.childnet.com/blog/what-is-a-snapchat-streak/ Accessed: 2022-07-21.

[29] Hyunsung Cho, DaEun Choi, Donghwi Kim, Wan Ju Kang, Eun Kyoung Choe, and Sung-Ju Lee. 2021. Reflect, Not Regret: Understanding Regretful Smartphone Use with App Feature-Level Analysis. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 456 (oct 2021), 36 pages. https://doi.org/10.1145/3479600 ✓.

[30] Robert B. Cialdini. 1987. *Influence: The Psychology of Persuasion.* Harper Business.

[31] Gregory Conti and Edward Sobiesk. 2010. Malicious Interface Design: Exploiting the User. In *Proceedings of the 19th International Conference on World Wide Web* (Raleigh, North Carolina, USA) *(WWW '10)*. Association for Computing Machinery, New York, NY, USA, 271–280. https://doi.org/10.1145/1772690.1772719 ✓.

[32] Thomas H. Davenport and John C. Beck. 2001. *Attention Economy: Understanding the New Currency of Business.* Harvard Business School Press.

[33] Linda Di Geronimo, Larissa Braz, Enrico Fregnan, Fabio Palomba, and Alberto Bacchelli. 2020. UI Dark Patterns and Where to Find Them: A Study on Mobile Applications and User Perception. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376600 ✓.

[34] Sarah Diefenbach and Kim Borrmann. 2019. The Smartphone as a Pacifier and Its Consequences: Young Adults' Smartphone Usage in Moments of Solitude and Correlations to Self-Reflection. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 1–14. ✓.

[35] D. Harold Doty and William H. Glick. 1994. Typologies As a Unique Form Of Theory Building: Toward Improved Understanding and Modeling. *Academy of Management Review* 19, 2 (1994), 230–251. https://doi.org/10.5465/amr.1994.9410210748

[36] Duolingo. 2022. Duolingo. The free, fun, and effective way to learn a language! https://www.duolingo.com/ Accessed: 2022-09-08.

[37] European Data Protection Board 2022. Guidelines 3/2022 on Dark patterns in social media platform interfaces: How to recognise and avoid them. https://edpb.europa.eu/our-work-tools/documents/public-consultations/2022/guidelines-32022-dark-patterns-social-media_en Accessed: 2022-09-11.

[38] Facebook. 2022. How do I permanently delete my Facebook account? https://www.facebook.com/help/224562897555674 Accessed: 2022-07-21.

[39] Samira Farivar, Fang Wang, and Ofir Turel. 2022. Followers' Problematic Engagement with Influencers on Social Media: An Attachment Theory Perspective. *Comput. Hum. Behav.* 133, C (aug 2022), 11 pages. https://doi.org/10.1016/j.chb.2022.107288 ✓.

[40] Federal Trade Commission 2021. Bringing Dark Patterns to Light: An FTC Workshop. https://www.ftc.gov/news-events/events/2021/04/bringing-dark-patterns-light-ftc-workshop Accessed: 2022-9-11.

[41] Dan Fitton and Janet C. Read. 2019. Creating a Framework to Support the Critical Consideration of Dark Design Aspects in Free-to-Play Apps. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children* (Boise, ID, USA) *(IDC '19)*. Association for Computing Machinery, New York, NY, USA, 407–418. https://doi.org/10.1145/3311927.3323136 ✓.

[42] Jonathan Safran Foer. 2016. Technology is diminishing us. https://www.theguardian.com/books/2016/dec/03/jonathan-safran-foer-technology-diminishing-us Accessed: 2021-08-06.

[43] T. Gilovich, V. H. Medvec, and K. Savitsky. 2000. The spotlight effect in social judgment: An egocentric bias in estimates of the salience of one's own actions and appearance. *Journal of Personality and Social Psychology* 78, 2 (2000), 211–22. https://doi.org/10.1037/0022-3514.78.2.211

[44] Colin M. Gray, Shruthi Sai Chivukula, and Ahreum Lee. 2020. What Kind of Work Do "Asshole Designers" Create? Describing Properties of Ethical Concern on Reddit. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 61–73. https://doi.org/10.1145/3357236.3395486 ✓.

[45] Colin M. Gray, Shruthi Sai Chivukula, Kassandra Melkey, and Rhea Manocha. 2021. Understanding "Dark" Design Roles in Computing Education. In *Proceedings of the 17th ACM Conference on International Computing Education Research* (Virtual Event, USA) *(ICER 2021)*. Association for Computing Machinery, New York, NY, USA, 225–238. https://doi.org/10.1145/3446871.3469754 ✓.

[46] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018. The Dark (Patterns) Side of UX Design. In *Proceedings of the 2018 CHI*

Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–14. ✓.

[47] Johanna Gunawan, Amogh Pradeep, David Choffnes, Woodrow Hartzog, and Christo Wilson. 2021. A Comparative Study of Dark Patterns Across Web and Mobile Modalities. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 377 (oct 2021), 29 pages. https://doi.org/10.1145/3479521 ✓.

[48] Ellie Harmon and Melissa Mazmanian. 2013. Stories of the Smartphone in Everyday Discourse: Conflict, Tension & Instability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) (CHI '13). ACM, New York, NY, USA, 1051–1060. https://doi.org/10.1145/2470654.2466134

[49] Joshua Harwood, Julian J. Dooley, Adrian J. Scott, and Richard Joiner. 2014. Constantly connected – The effects of smart-devices on mental health. *Computers in Human Behavior* 34 (2014), 267–272. https://doi.org/10.1016/j.chb.2014.02.006 ✓.

[50] Md Rajibul Hasan, Ashish Kumar Jha, and Yi Liu. 2018. Excessive Use of Online Video Streaming Services. *Comput. Hum. Behav.* 80, C (mar 2018), 220–228. https://doi.org/10.1016/j.chb.2017.11.020 ✓.

[51] Ralph Hertwig and Till Grüne-Yanoff. 2017. Nudging and Boosting: Steering or Empowering Good Decisions. *Perspectives on Psychological Science* 12 (08 2017), 174569161770249. https://doi.org/10.1177/1745691617702496

[52] Min-Wei Hung, Chien Wen (Tina) Yuan, Nanyi Bi, Yi-Chao Chen, Wan-Chen Lee, Ming-Chyi Huang, and Chuang-Wen You. 2022. To Use or Abuse: Opportunities and Difficulties in the Use of Multi-Channel Support to Reduce Technology Abuse by Adolescents. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 125 (apr 2022), 27 pages. https://doi.org/10.1145/3512972 ✓.

[53] Se-Hoon Jeong, HyoungJee Kim, Jung-Yoon Yum, and Yoori Hwang. 2016. What type of content are smartphone users addicted to?: SNS vs. games. *Computers in Human Behavior* 54 (2016), 10 – 17. https://doi.org/10.1016/j.chb.2015.07.035

[54] Lexie Kane. 2019. The Attention Economy. https://www.nngroup.com/articles/attention-economy/ Accessed: 2021-12-27.

[55] M J Kelly. 2021. What are deceptive design patterns and how can you spot them? https://blog.mozilla.org/en/internet-culture/mozilla-explains/deceptive-design-patterns/ Accessed: 2022-8-28.

[56] Dongil Kim, JeeEun Karin Nam, JungSu Oh, and Min Chul Kang. 2016. A latent profile analysis of the interplay between PC and smartphone in problematic internet use. *Computers in Human Behavior* 56 (2016), 360–368. https://doi.org/10.1016/j.chb.2015.11.009

[57] Dong Hoo Kim, Natalee Kate Seely, and Jong-Hyuok Jung. 2017. Do You Prefer, Pinterest or Instagram? The Role of Image-Sharing SNSs and Self-Monitoring in Enhancing Ad Effectiveness. *Comput. Hum. Behav.* 70, C (may 2017), 535–543. https://doi.org/10.1016/j.chb.2017.01.022 ✓.

[58] Kagan Kircaburun, Saleem Alhabash, Şule Betül Tosuntaş, and Mark D Griffiths. 2020. Uses and gratifications of problematic social media use among university students: A simultaneous examination of the Big Five of personality traits, social media platforms, and social media use motives. *International journal of mental health and addiction* 18 (2020), 525–547. https://doi.org/10.1007/s11469-018-9940-6

[59] Minsam Ko, Seungwoo Choi, Koji Yatani, and Uichin Lee. 2016. Lock n' LoL: Group-Based Limiting Assistance App to Mitigate Smartphone Distractions in Group Activities. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 998–1010. https://doi.org/10.1145/2858036.2858568

[60] Minsam Ko, Subin Yang, Joonwon Lee, Christian Heizmann, Jinyoung Jeong, Uichin Lee, Daehee Shin, Koji Yatani, Junehwa Song, and Kyong-Mee Chung. 2015. NUGU: A Group-Based Intervention App for Improving Self-Regulation of Limiting Smartphone Use. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) (CSCW '15). Association for Computing Machinery, New York, NY, USA, 1235–1245. https://doi.org/10.1145/2675133.2675244

[61] Konrad Kollnig, Siddhartha Datta, and Max Van Kleek. 2021. I Want My App That Way: Reclaiming Sovereignty Over Personal Devices. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 393, 8 pages. ✓.

[62] Logan Kugler. 2020. Are We Addicted to Technology? *Commun. ACM* 63, 8 (July 2020), 15–16. https://doi.org/10.1145/3403966

[63] Klodiana Lanaj, Russell E. Johnson, and Christopher M. Barnes. 2014. Beginning the workday yet already depleted? Consequences of late-night smartphone use and sleep. *Organizational Behavior and Human Decision Processes* 124, 1 (2014), 11 – 23. https://doi.org/10.1016/j.obhdp.2014.01.001

[64] Simone Lanette, Phoebe K. Chua, Gillian Hayes, and Melissa Mazmanian. 2018. How Much is 'Too Much'?: The Role of a Smartphone Addiction Narrative in Individuals' Experience of Use. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW, Article 101 (Nov. 2018), 22 pages. https://doi.org/10.1145/3274370

[65] Stephen M. Lebruto, William J. Quain, and A. A. Ashley. 1995. Menu Engineering: A Model Including Labor. *Hospitality Review* 13 (1995), 41–49.

[66] Uichin Lee, Joonwon Lee, Minsam Ko, Changhun Lee, Yuhwan Kim, Subin Yang, Koji Yatani, Gahgene Gweon, Kyong-Mee Chung, and Junehwa Song.

2014. Hooked on Smartphones: An Exploratory Study on Smartphone Overuse among College Students. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (CHI '14). Association for Computing Machinery, New York, NY, USA, 2327–2336. https://doi.org/10.1145/2556288.2557366 ✓.

[67] Chris Lewis. 2014. Temporal Dark Patterns. In *Irresistible Apps: Motivational Design Patterns for Apps, Games, and Web-based Communities*, Chris Lewis (Ed.). Apress, Berkeley, CA, 103–110. https://doi.org/10.1007/978-1-4302-6422-4_9

[68] Paul Lewis. 2017. 'Our minds can be hijacked': the tech insiders who fear a smartphone dystopia. https://www.theguardian.com/technology/2017/oct/05/smartphone-addiction-silicon-valley-dystopia Accessed: 2021-12-25.

[69] Alessandro Liberati, Douglas G. Altman, Jennifer Tetzlaff, Cynthia Mulrow, Peter C. Gøtzsche, John P. A. Ioannidis, Mike Clarke, P. J. Devereaux, Jos Kleijnen, and David Moher. 2009. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLOS Medicine* 6, 7 (07 2009), 1–28. https://doi.org/10.1371/journal.pmed.1000100

[70] Jamie Luguri and Lior Strahilevitz. 2019. Shining a Light on Dark Patterns. (Aug. 2019). https://doi.org/10.2139/ssrn.3431205

[71] Kai Lukoff. 2022. *Designing to Support Sense of Agency for Time Spent on Digital Interfaces*. Ph. D. Dissertation. University of Washington.

[72] Kai Lukoff, Alexis Hiniker, Colin M. Gray, Arunesh Mathur, and Shruthi Sai Chivukula. 2021. What Can CHI Do About Dark Patterns?. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI EA '21). Association for Computing Machinery, New York, NY, USA, Article 102, 6 pages. https://doi.org/10.1145/3411763.3441360

[73] Kai Lukoff, Ulrik Lyngs, Himanshu Zade, J. Vera Liao, James Choi, Kaiyue Fan, Sean A. Munson, and Alexis Hiniker. 2021. How the Design of YouTube Influences User Sense of Agency. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 368, 17 pages. https://doi.org/10.1145/3411764.3445467 ✓.

[74] Kai Lukoff and Sean Munson. 2019. Digital wellbeing is way more than just reducing screen time. https://uxdesign.cc/digital-wellbeing-more-than-just-reducing-screen-time-46223db9f057. https://uxdesign.cc/digital-wellbeing-more-than-just-reducing-screen-time-46223db9f057 Accessed: 2020-10-1.

[75] Kai Lukoff, Cissy Yu, Julie Kientz, and Alexis Hiniker. 2018. What Makes Smartphone Use Meaningful or Meaningless? *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 22 (mar 2018), 26 pages. https://doi.org/10.1145/3191754 ✓.

[76] Ulrik Lyngs, Kai Lukoff, Petr Slovak, Reuben Binns, Adam Slack, Michael Inzlicht, Max Van Kleek, and Nigel Shadbolt. 2019. Self-Control in Cyberspace: Applying Dual Systems Theory to a Review of Digital Self-Control Tools. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–18.

[77] Ulrik Lyngs, Kai Lukoff, Petr Slovak, William Seymour, Helena Webb, Marina Jirotka, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2020. 'I Just Want to Hack Myself to Not Get Distracted': Evaluating Design Interventions for Self-Control on Facebook. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376672 ✓.

[78] Aditi M. Bhoot, Mayuri A. Shinde, and Wricha P. Mishra. 2020. 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* (Online, India) (IndiaHCI 2020). Association for Computing Machinery, New York, NY, USA, 24–33. https://doi.org/10.1145/3429290.3429293 ✓.

[79] Gloria Mark, Shamsi Iqbal, Mary Czerwinski, and Paul Johns. 2015. Focused, Aroused, but So Distractible: Temporal Perspectives on Multitasking and Communications. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) (CSCW '15). ACM, New York, NY, USA, 903–916. https://doi.org/10.1145/2675133.2675221

[80] Gloria Mark, Yiran Wang, and Melissa Niiya. 2014. Stress and Multitasking in Everyday College Life: An Empirical Study of Online Activity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (CHI '14). ACM, New York, NY, USA, 41–50. https://doi.org/10.1145/2556288.2557361

[81] Arunesh Mathur, Gunes Acar, Michael J. Friedman, Elena Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan. 2019. Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 81 (nov 2019), 32 pages. https://doi.org/10.1145/3359183 ✓.

[82] Arunesh Mathur, Mihir Kshirsagar, and Jonathan Mayer. 2021. What Makes a Dark Pattern... Dark? Design Attributes, Normative Considerations, and Measurement Methods. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 360, 18 pages. ✓.

[83] Lee McGuigan, Helen Nissenbaum, Beate Roessler, and Daniel Susser. 2020. Online Manipulation Workshop. https://www.dli.tech.cornell.edu/manipulation Accessed: 2022-9-11.

[84] Dar Meshi, Carmen Morawetz, and Hauke Heekeren. 2013. Nucleus accumbens response to gains in reputation for the self relative to gains for others predicts social media use. *Frontiers in Human Neuroscience* 7 (2013), 439. https://doi.org/10.3389/fnhum.2013.00439 ✓.

[85] Thomas Mildner and Gian-Luca Savino. 2021. Ethical User Interfaces: Exploring the Effects of Dark Patterns on Facebook. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 464, 7 pages. ✓.

[86] David Moher, Alessandroi Liberati, Jennifer Tetzlaff, and Douglas G. Altman. 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine* 151, 4 (2009), 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135 PMID: 19622511.

[87] Alberto Monge Roffarello and Luigi De Russis. 2019. The Race Towards Digital Wellbeing: Issues and Opportunities. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300616

[88] Sara Morrison. 2021. Dark patterns, the tricks websites use to make you say yes, explained. https://www.vox.com/recode/22351108/dark-patterns-ui-web-design-privacy Accessed: 2022-09-11.

[89] Arvind Narayanan, Arunesh Mathur, Marshini Chetty, and Mihir Kshirsagar. 2020. Dark Patterns: Past, Present, and Future. *Commun. ACM* 63, 9 (Aug. 2020), 42–47. https://doi.org/10.1145/3397884

[90] Denver Nicks. 2015. LinkedIn to Pay $13 Million in Spam Settlement. *Time* (Oct. 2015). https://time.com/4062519/linkedn-spam-settlement/

[91] Pawarat Nontasil and Stephen J. Payne. 2019. Emotional Utility and Recall of the Facebook News Feed. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–9. ✓.

[92] Anahad O'Connor. 2019. Sugary Drink Ban Tied to Health Improvements at Medical Center. *The New York Times* (Oct. 2019). https://www.nytimes.com/2019/10/28/well/eat/sugary-drink-soda-ban-health-medical-center.html

[93] Ihudiya Finda Ogbonnaya-Ogburu, Angela D R Smith, Alexandra To, and Kentaro Toyama. 2020. Critical Race Theory for HCI. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3313831.3376392

[94] Antti Oulasvirta, Tye Rattenbury, Lingyi Ma, and Eeva Raita. 2012. Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing* 16, 1 (Jan 2012), 105–114. https://doi.org/10.1007/s00779-011-0412-2

[95] Chunjong Park, Junsung Lim, Juho Kim, Sung-Ju Lee, and Dongman Lee. 2017. Don't Bother Me. I'm Socializing! A Breakpoint-Based Smartphone Notification System. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 541–554. https://doi.org/10.1145/2998181.2998189

[96] Joonyoung Park, Hyunsoo Lee, Sangkeun Park, Kyong-Mee Chung, and Uichin Lee. 2021. GoldenTime: Exploring System-Driven Timeboxing and Micro-Financial Incentives for Self-Regulated Phone Use. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 702, 17 pages. ✓.

[97] Martin Pielot, Karen Church, and Rodrigo de Oliveira. 2014. An In-Situ Study of Mobile Phone Notifications. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services* (Toronto, ON, Canada) *(MobileHCI '14)*. Association for Computing Machinery, New York, NY, USA, 233–242. https://doi.org/10.1145/2628363.2628364

[98] Charlie Pinder. 2017. The Anti-Influence Engine: Escaping the Diabolical Machine of Pervasive Advertising. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI EA '17)*. Association for Computing Machinery, New York, NY, USA, 770–781. https://doi.org/10.1145/3027063.3052762 ✓.

[99] Charlie Pinder, Jo Vermeulen, Benjamin R. Cowan, and Russell Beale. 2018. Digital Behaviour Change Interventions to Break and Form Habits. *ACM Transactions on Computer-Human Interaction* 25, 3, Article 15 (June 2018), 66 pages. https://doi.org/10.1145/3196830

[100] Aarathi Prasad, Lucas S. LaFreniere, Vaasu Taneja, and Zoe Beals. 2021. Addressing Problematic Smartphone Use with a Personalized, Goal-Based Approach. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers* (Virtual, USA) *(UbiComp '21)*. Association for Computing Machinery, New York, NY, USA, 131–134. https://doi.org/10.1145/3460418.3479319

[101] Mashfiqui Rabbi, Min Hane Aung, Mi Zhang, and Tanzeem Choudhury. 2015. MyBehavior: Automatic Personalized Health Feedback from User Behaviors and Preferences Using Smartphones. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 707–718. https://doi.org/10.1145/2750858.2805840

[102] Eric Ravenscraft. 2020. How to Spot—and Avoid—Dark Patterns on the Web. https://www.wired.com/story/how-to-spot-avoid-dark-patterns/ Accessed: 2022-09-11.

[103] Eric Reicin. 2021. Understanding Dark Patterns: How To Stay Out Of The Gray Areas. https://www.forbes.com/sites/forbesnonprofitcouncil/2021/04/19/understanding-dark-patterns-how-to-stay-out-of-the-gray-areas/ Accessed: 2022-9-11.

[104] Alberto Monge Roffarello and Luigi De Russis. 2021. Understanding, Discovering, and Mitigating Habitual Smartphone Use in Young Adults. *ACM Trans. Interact. Intell. Syst.* 11, 2, Article 13 (jul 2021), 34 pages. https://doi.org/10.1145/3447991 ✓.

[105] Colleen Saffrey, Amy Summerville, and Neal Roese. 2008. Praise for regret: People value regret above other negative emotions. *Motivation and emotion* 32 (04 2008), 46–54. https://doi.org/10.1007/s11031-008-9082-4

[106] Christine Satchell and Paul Dourish. 2009. Beyond the User: Use and Non-Use in HCI. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7* (Melbourne, Australia) *(OZCHI '09)*. Association for Computing Machinery, New York, NY, USA, 9–16. https://doi.org/10.1145/1738826.1738829

[107] Brennan Schaffner, Neha Lingareddy, and Marshini Chetty. 2022. Understanding Account Deletion and Relevant Dark Patterns On Social Media [Pre-print]. *Computer Supported Cooperative Work (CSCW)* (2022). https://airlab.cs.uchicago.edu/files/2022/06/PREPRINT_Understanding_Account_Deletion_CSCW2022-1.pdf

[108] Caroline Sinders. 2022. What's In a Name? https://medium.com/@carolinesinders/whats-in-a-name-unpacking-dark-patterns-versus-deceptive-design-e96068627ec4 Accessed: 2022-8-28.

[109] Manya Sleeper, Alessandro Acquisti, Lorrie Faith Cranor, Patrick Gage Kelley, Sean A. Munson, and Norman Sadeh. 2015. I Would Like To..., I Shouldn't..., I Wish I...: Exploring Behavior-Change Goals for Social Networking Sites. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 1058–1069. https://doi.org/10.1145/2675133.2675193

[110] Aaron Smith, Skye Toor, and Patrick Van Kessel. 2018. Many Turn to YouTube for Children's Content, News, How-To Lessons. https://www.pewresearch.org/internet/2018/11/07/many-turn-to-youtube-for-childrens-content-news-how-to-lessons/ Accessed: 2020-3-3.

[111] Adrienne So. 2019. Instagram Will Test Hiding 'Likes' in the US Starting Next Week. https://www.wired.com/story/instagram-hiding-likes-adam-mosseri-tracee-ellis-ross-wired25/ Accessed: 2021-08-06.

[112] Than Htut Soe, Oda Elise Nordberg, Frode Guribye, and Marija Slavkovik. 2020. Circumvention by Design - Dark Patterns in Cookie Consent for Online News Outlets. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (Tallinn, Estonia) *(NordiCHI '20)*. Association for Computing Machinery, New York, NY, USA, Article 19, 12 pages. https://doi.org/10.1145/3419249.3420132

[113] J. E. R. Staddon and D. T. Cerutti. 2003. Operant Conditioning. *Annual Review of Psychology* 54, 1 (2003), 115–144. https://doi.org/10.1146/annurev.psych.54.101601.145124

[114] Daniel Susser and Vincent Grimaldi. 2021. Measuring Automated Influence: Between Empirical Evidence and Ethical Values. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (Virtual Event, USA) *(AIES '21)*. Association for Computing Machinery, New York, NY, USA, 242–253. https://doi.org/10.1145/3461702.3462532 ✓.

[115] Matthis Synofzik, Gottfried Vosgerau, and Albert Newen. 2008. Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition* 17, 1 (2008), 219–239. https://doi.org/10.1016/j.concog.2007.03.010

[116] Richard H. Thaler and Cass R. Sunstein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness.* Yale University Press, New Haven, CT, USA.

[117] The National Law Review 2022. "Dark Patterns" Are Focus of Regulatory Scrutiny in the United States and Europe. https://www.natlawreview.com/article/dark-patterns-are-focus-regulatory-scrutiny-united-states-and-europe Accessed: 2022-9-11.

[118] Anja Thieme, Danielle Belgrave, and Gavin Doherty. 2020. Machine Learning in Mental Health: A Systematic Review of the HCI Literature to Support the Development of Effective and Implementable ML Systems. *ACM Transactions on Computer-Human Interaction* 27, 5, Article 34 (Aug. 2020), 53 pages. https://doi.org/10.1145/3398069

[119] Jonathan A. Tran, Katie S. Yang, Katie Davis, and Alexis Hiniker. 2019. Modeling the Engagement-Disengagement Cycle of Compulsive Phone Use. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. ✓.

[120] Sherry Turkle. 2011. *Alone Together: Why We Expect More from Technology and Less from Each Other.* Basic Books, Inc., New York, NY, USA.

[121] Amos Tversky and Daniel Kahneman. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science* 185, 4157 (1974), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

[122] Twitter. 2021. About Verified Accounts. https://help.twitter.com/en/managing-your-account/about-twitter-verified-accounts Accessed: 2021-12-27.

[123] Unhook. 2022. Unhook - Remove YouTube Recommended Videos. https://unhook.app/ Accessed: 2022-07-20.

[124] Northwestern University. 2021. Inclusive Language in Technology: Information Technology - Northwestern University. https://www.it.northwestern.edu/about/it-projects/dei/glossary.html Accessed: 2022-9-12.

[125] Bahromjon Urmanov and Shin Hoyoung. 2021. An Empirical Investigation on the Factors Affecting Smartphone Addiction in Digital Era: Effect of Social Networking Services and Instant Messaging Applications Usage. In *The 5th International Conference on Future Networks & Distributed Systems* (Dubai, United Arab Emirates) *(ICFNDS 2021)*. Association for Computing Machinery, New York, NY, USA, 473–484. https://doi.org/10.1145/3508072.3508169 ✓.

[126] Philippe Verduyn, David Lee, Jiyoung Park, Holly Shablack, Ariana Orvell, Joseph Bayer, Oscar Ybarra, John Jonides, and Ethan Kross. 2015. Passive Facebook Usage Undermines Affective Well-Being: Experimental and Longitudinal Evidence. *Journal of Experimental Psychology General* 144, 2 (02 2015), 480–488. https://doi.org/10.1037/xge0000057

[127] Fiona Westin and Sonia Chiasson. 2021. "It's So Difficult to Sever That Connection": The Role of FoMO in Users' Reluctant Privacy Behaviours. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 550, 15 pages. https://doi.org/10.1145/3411764.3445104

[128] Kelly Widdicks, Daniel Pargman, and Staffan Bjork. 2020. Backfiring and Favouring: How Design Processes in HCI Lead to Anti-Patterns and Repentant Designers. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*. Association for Computing Machinery, New York, NY, USA, Article 16, 12 pages. ✓.

[129] World Wide Web Foundation 2022. Deceptive Design. https://techlab.webfoundation.org/deceptive-design/overview Accessed: 2022-8-29.

[130] José P. Zagal, Staffan Björk, and Chris Lewis. 2013. Dark patterns in the design of games. In *Proceedings of the 8th International Conference on the Foundations of Digital Games, FDG 2013* (Chania, Crete, Greece), Georgios N. Yannakakis, Espen Aarseth, Kristine Jørgensen, and James C. Lester (Eds.). Society for the Advancement of the Science of Digital Games, Santa Cruz, CA, USA, 39–46. ✓.

[131] Eric Zeng, Tadayoshi Kohno, and Franziska Roesner. 2021. What Makes a "Bad" Ad? User Perceptions of Problematic Online Advertising. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 361, 24 pages. ✓.

[132] Mingrui Ray Zhang, Kai Lukoff, Raveena Rao, Amanda Baughan, and Alexis Hiniker. 2022. Monitoring Screen Time or Redesigning It? Two Approaches to Supporting Intentional Social Media Use. In *CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 60, 19 pages. https://doi.org/10.1145/3491102.3517722 ✓.