If Deceptive Patterns are the problem, are Fair Patterns the solution?

Tim de Jonge
Radboud University
Nijmegen, Netherlands
tim.dejonge@ru.nl

Hanna Schraffenberger
Radboud University
Nijmegen, Netherlands
hanna.schraffenberger@ru.nl

Jorrit Geels
Radboud University
Nijmegen, Netherlands
jorrit.geels@me.com

Jaap-Henk Hoepman
Radboud University
Nijmegen, Netherlands
jhh@cs.ru.nl

Marie-Sophie Simon
Radboud University
Nijmegen, Netherlands
marie-sophie.simon@ru.nl

Frederik Zuiderveen Borgesius
Radboud University
Nijmegen, Netherlands
frederikzb@cs.ru.nl

Abstract

Researchers and legislators increasingly worry about *deceptive patterns*: common tricks on websites and in apps that make users do things they did not intend to do (previously: dark patterns). If these deceptive patterns are a problem, could *"fair patterns"* be the solution? We highlight several caveats to this approach. First, it is not obvious what it means for a design pattern to be *fair*. What is fair depends on the context and even within the same context, people disagree on what fairness means. Moreover, one fair design element does not guarantee a fair overall design. Combining these objections, it may be inappropriate to call a design pattern fair. Second, not all problems are adequately addressed by interventions at the design level. If all possible choices are unfair, design alone cannot make the situation fair. Societal problems must be solved at a societal scale, although design can contribute through incremental improvements. Progress in interface design does not need the concept of fairness: empirically informed solutions for specific problems appear more practical.

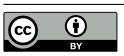
CCS Concepts

• **Human-centered computing → HCI theory, concepts and models**; **Interaction design theory, concepts and paradigms**; *User interface design.*

Keywords

Interaction Design, Fairness, Dark Patterns, Deceptive Patterns

ACM Reference Format:
Tim de Jonge, Hanna Schraffenberger, Jorrit Geels, Jaap-Henk Hoepman, Marie-Sophie Simon, and Frederik Zuiderveen Borgesius. 2025. If Deceptive Patterns are the problem, are Fair Patterns the solution?. In *The 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25), June 23–26, 2025, Athens, Greece.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3715275.3732199



FAccT '25, Athens, Greece

ACM ISBN 979-8-4007-1482-5/25/06
https://doi.org/10.1145/3715275.3732199

1 Introduction

Fairness is increasingly receiving attention across research fields: in law [32], whether it be data protection law [19, 49], competition law [31], or consumer protection law [86]; in computer science, whether in search [84], human recommendation [48], or healthcare [17]; or overarching these disciplines [46, 82, 85]. In this work, we focus on fairness in a different domain: user interaction design. Here, the interest in fairness has lead to the study of *deceptive patterns*. While various definitions exist (see Mathur et al. [58] for an overview), deceptive patterns are commonly understood as tricks on websites and in apps that make users do things they did not intend to do [9]. More specifically, they refer to "user interfaces whose designers knowingly confuse users, make it difficult for users to express their actual preferences, or manipulate users into taking certain actions" [52, p. 43]. In other words, a deceptive pattern is an interaction design choice that steers users towards taking actions that are not in line with their true interest and, instead, benefit an online service [57].

Deceptive patterns are widely used in digital products, and can be found on social media [60], e-commerce [57], and in games [87]. Concrete examples range from popups that push users to accept cookies [36] and unnecessarily difficult newspaper cancellation processes [71], to fake countdown timers on webshops that urge users to order products [80]. As discussed by Mathur et al. [58], deceptive patterns pose several potential harms, ranging from individual harms (e.g., invasion of privacy, financial loss, cognitive burdens) to collective harms (e.g., loss of market trust).

Chamorro et al. [13] find that designers can use these potentially manipulative patterns without intending to. Practitioners have also expressed the need for practical guidance [4]. As deceptive patterns are a widely acknowledged problem, practical guidance on how to avoid these deceptive patterns should be common.

Although we have well-documented examples of bad practices to avoid [e.g., 35, 58], there does not appear to be sufficient research effort towards positive examples to follow, or towards advice aimed at practitioners. From an HCI perspective, it is clear that many traditional design guidelines (such as the widely used ten heuristics by Nielsen [61]) with their emphasis on usability provide insufficient guidance, as "a usable application does not imply an ethical one" [23, p.2]. Research aimed at fair design has mostly resulted in high-level design principles, rather than actionable advice [2, 5, 59, 64, 73].

We can contrast this development of fairness as a design value with how privacy has come to prominence. While fair design currently remains abstract and theoretical, privacy by design has gone through a rich discussion between theory and practice. In the early days of the internet, Chaum and colleagues identified that privacy could come under fire in the digital age, and pioneered some initial interventions to ensure digital privacy through technical means [e.g., 14–16]. Ever since, legal scholars, practitioners with an interest in privacy, and those who are more theoretically inclined have iterated on an understanding of privacy [11]. From a practical perspective, more Privacy-Enhancing Technologies were invented [39]. Privacy-by-Design bound these technologies together under a common umbrella to address privacy from the start of the design process [12, 43]. This high-level perspective sparked the need for more actionable advice, which inspired *privacy patterns*; reusable solutions to the privacy problems that arise in many designs [25, 38, 51]. Insights into privacy progressed as lawmakers started protecting privacy in the digital realm, and *privacy strategies* translated these legal requirements into practice [20, 40]. At the end of this development, we can now see privacy as a central value both in digital legislation and in system design.

In the development of privacy as design value, we can see a back-and-forth between theoretical and practical insights. In the development of fairness as design value, the practical insights appear to be lagging behind. By analogy with privacy design patterns, one would expect there to be *fair patterns*; reusable solutions that would improve the fairness of a given design. To an extent, researchers are trying to create such patterns [63] — but efforts are still in their infancy.

We are an interdisciplinary team of computer scientists, designers, and legal scholars, combining insights from disciplines to investigate the difficulties of realizing fair patterns. Our contributions are twofold: First, we investigate the potential of fair patterns as solution to deceptive patterns. Second, we analyze *fairness* as a constructive term, used in a new domain. To this end, we explore two research questions. Our first research question is: What is fairness in the context of design patterns? Our second research question is: How can we apply fairness to design patterns?

In Section 2, we introduce fair patterns through discussing the history of deceptive patterns, and design patterns in general. In Section 3, we explore what the meaning of fairness in the context of design patterns. In Section 4, we investigate the limitations of fair patterns. In Section 5, we provide an outlook on how to provide practical guidance to practitioners and outline potential research avenues. We conclude in Section 6.

2 Background

Before we dive into the possibility of fair patterns, we highlight the history of deceptive patterns and design patterns in general. Design patterns are repeatable solutions to recurring design problems [62]. Alexander et al. [3] originally introduced design patterns as a way to structure knowledge found in the architectural domain [3]. These design patterns where presented in a structured fashion, but conceptually, design patterns ranged very broadly. The book begins as broad as *the countryside* as a high-level pattern [3, p. 36], but gets as specific as a *waist-high shelf* [3, p. 922]. Later, in 1994, Gamma et al. [30] popularized design patterns in computer science. Ever since, computer scientists have used design patterns, categorizing them either by technology [55], domain [79], or specific application [42].

In this case, we are restricting ourselves to design patterns as they relate to User Interaction (UI) design. Here, one prominent class of design patterns is called *deceptive patterns* [9]. Originally, Brignull had designated these patterns as *dark patterns*. In recent years, the common name for these patterns has been shifting from *dark patterns* to *deceptive patterns*. There are two main reasons for this change. First, the term *dark patterns* is insufficiently descriptive. The term appeals to a problem with these patterns, but the vagueness of the term has led to scattered interpretations, and confusion among practitioners [13, 58]. Second, the use of the word *dark* to mean *bad* carries racial connotations. As such, we can use other terms to be more inclusive [26].

Deceptive patterns are heavily researched and discussed in the HCI community [e.g., 34], which has identified and classified various concrete types of deceptive patterns (e.g., hidden costs, sneak into basket, trick questions) [e.g., 9] and organized them in deceptive patterns taxonomies [e.g., 52] and ontologies [e.g., 35]. Likewise, strategies commonly used to modify user's choice architectures have been mapped out [e.g., 33]. Mathur et al. [58] provides a comprehensive overview of problematic approaches such as inducing false beliefs, withholding relevant information, removing choices that should be available, disadvantageous treatment of specific user groups, or hiding influential mechanisms. Broadly, Deceptive Patterns make choices that benefit the business more easily accessible than choices that benefit the user [58].

Considering these ways in which deceptive patterns modify users' choice architecture and the range of potential harms, it is not surprising that deceptive patterns are commonly considered problematic, unfair, or unethical [18, 21, 34] — i.e., the opposite of what a well-meaning designer would aspire to.

We are not the first to imagine fair patterns as a potential solution to deceptive patterns. In the beginning of 2024, an interdisciplinary workshop brought together researchers from academia, businesses and enforcement with legal, human-computer interaction and economic backgrounds, to jointly develop the concept of fair patterns [68], leading some attendants to question whether fair patterns exist [41]. Further, Potel-Saville and Rocha [63] presented a taxonomy of fair patterns, proposing solutions such as "protective defaults" as an alternative to "harmful defaults" and "plain and empowering language" as an alternative to "misleading or obstructing Language". In addition to such research efforts, the commercial site fairpatterns.com [1] provides resources about fair patterns for businesses, regulators and law firms, among which a fair patterns Library that illustrates the taxonomy by Potel-Saville and Rocha [63].

Others have proposed fairer responses to deceptive patterns. These include so-called *"bright patterns"* [36, 65, 66], which nudge users towards the choice that likely is in *their* best interest and *"reflective patterns"* [e.g., 77], which foster reflective and deliberate decision-making through designing-in friction. In this work, we will address the difficulties fair patterns encounter, and sketch a potential path forward.

3 What is fairness in the context of design patterns?

In this section, we will explore how a design pattern can be fair. A logical step towards this goal could be to precisely define fairness. The term "fair" is already in use, across several related fields. Generally, authors use fairness rather broadly [e.g., 19, 32, 86]. Gerard [31] specifically remarks on the broadness of the fairness in competition law. Even within just the General Data Protection Regulation, the meaning of fairness can be difficult to pin down [53]. In algorithmic design, fairness can narrowly refer to an absence of discrimination, or broadly, to a morally just algorithm [46, 85]. As our argumentation does not depend on a particular definition of fairness, we choose to interpret "fair" as morally just, in some broad sense. This follows use of the term fair in both legal doctrine, and casual speech [31, 86]. We emphasize that the difficulties around defining fairness and the broadness of the term pose a challenge for countering deceptive design with fair patterns.

3.1 Fairness is contextual

The design-pattern level appears too narrow to apply the label "fair': whether a design is fair depends on the context in which it is deployed. Selbst et al. [70] defines the *"portability trap"*: "Failure to understand how repurposing algorithmic solutions designed for one social context may be misleading, inaccurate, or otherwise do harm when applied to a different context" [70]. Although Selbst et al. [70] define this notion in the sphere of algorithmic fairness, the same argument holds in interface design. If we try to apply the concept of fairness to a design pattern, we assume that this design pattern is fair within a variety of design contexts. This is not obviously true.

As an example, If a user only intends to use a streaming service once, requiring a subscription that automatically renews seems unfair. Yet, automatic subscription renewal seems very sensible when discussing connection to the water main. Similarly, one may evaluate a design pattern used for self-enrichment differently than that same design pattern used to diminish inequality. Because of this, moral analysis of design will have to be done on a case-by-case basis [28]. In the related field of persuasive system design, Tannenbaum et al. [76] find that people accept nudges much more readily if the nudge aims to achieve something the person agrees with [76]. This complicates the discussion about the moral value of nudges per se: rather than discussing whether a design practice is acceptable, the discussion could quickly revolve around the moral standing of the underlying aim. Nudges are only a sub-category of design, but it seems reasonable to infer that this argument holds for design in a broader sense: for moral analysis, we must differentiate between the design practice, and the underlying aim of the design practice.

Fairness differs from other values, in the sense that fairness cannot be evaluated from just the technological intervention. For contrast, consider anonymity as design value. It is not very anonymous if a service requires the user to provide identification whenever they interact with the service. It is somewhat more anonymous if a service requires a user to register with an email and password, but requires no further identification. It is yet more anonymous if the service does not require any login to interact with a service. These statements are true regardless of the context in which these services operate; the level of anonymity can be evaluated from the technical intervention. The fairness of any of these interventions differs heavily by context; it seems fair when the government requires identification for a citizen to file their taxes, whereas it seems disproportionate for an app to demand identification for a casual game of chess. The same intervention can be fair or unfair, dependent on context.

Fairness is sensitive not only to change of application, but also in time. Currently, Dutch citizens are registered by default as organ donors. Before 2020, this was not the case; a Dutch citizen had to actively register if they wanted to become an organ donor. Here, we are dealing with default choices: bad defaults are commonly named as deceptive patterns [e.g., 11, 34, 57] and good defaults are suggested as fair patterns [63]. As the Netherlands became more secular, and support for donation-by-default rose, organ donation as a default has shifted from bad default to good default, without the design itself changing.

We can see that whether a design pattern is fair depends on the context in which it is deployed. One cannot evaluate the fairness of a design pattern without knowing sufficient context — be it sector, target audience, or country of deployment.

3.2 Fairness is contestable

Even within a particular context, fairness is a personal and political notion. We might be able to agree on broad principles of fairness. Aristotle posited that justice means treating like cases alike and different cases differently. Despite one's personal and political preferences, one might be able to agree with this notion. When we apply this notion to concrete problems, difficulties arise: which cases are alike, to what extent, and what does that mean for how we treat the case [8, 56]? In an abstract sense, we might be able to agree on what is fair. In a practical circumstance, whether something is fair depends on who defines what fair means.

Furthermore, fairness is a multifaceted notion [37]. Different authors stress different values as requirements for ethical design. For example, Abascal and Nicolle [2] list accessibility, privacy, autonomy, and economic factors as values to consider. This list is not beyond reproach — for example, transparency has been suggested as additional factor [5]. Even taking this list for granted, reasonable people can disagree on the amount of privacy one would be willing to give up to gain a measure of autonomy. It is hard to establish a design pattern as fair, when it is difficult to agree on what fair means in its own right. As a commonly used example, consider privacy notices [44]. Typically, privacy notices are written in legal jargon, or 'legalese'. In a strict sense, those providing the privacy notices are transparent: all information is present. However, the sheer amount of information is likely to overwhelm users. It may be less transparent to provide less information, but it would be more accessible. Improving a design in one regard can incur costs in a different regard. Even if we keep context constant, different notions of fairness exist, and as such, opinions on the fairness of a design can differ [72].

Again, fairness is not the first field to run into these problems, and we may be able to draw from the field of Privacy research. In 2005, Solove created a taxonomy of privacy, to clarify conceptual

challenges to the field [74]. This taxonomy presents privacy violations as a group of related harms. These privacy-related harms do not share a core characteristic between all of them, yet, they share important similarities, and as such, should be kept together under the umbrella of Privacy [74, page 562]. Similarly, we do not have to resolve Fair Design to a single, dense concept. We can view fairness as a set of related goals (e.g. privacy, accessibility); any work that works towards achieving these goals can be placed under the name of fairness. Our recommendation is to be specific about the harms addressed by a particular solution, and to accept that we may not agree on whether a given solution is fair [7].

In a positive phrasing, fairness is often about the appropriate balance of these different values in a given situation. This definition remains open to interpretation of what an appropriate balancing would be, and which interests one would need to balance. Still, it clarifies that we cannot create a fair design without significant room for adaptation to context and that fairness requires taking a moral stance.

4 How can we apply fairness to design patterns?

Previously, we have outlined the difficulties in defining "fairness". In the following section, we assume for the sake of argument that we have agreed on a definition of fairness, and instead, we shift our focus to applying the notion of fairness to design patterns. Fair design may not always be sufficient as a solution.

4.1 Design alone cannot solve unfair situations

In this paper, we scope design to mean user interaction design, which means there are many parts of the choice architecture that fair design cannot touch. Even if one would design a user interaction using only fair patterns in an appropriate manner, it is not guaranteed that this makes the resulting system fair. As an example, a vendor of concert tickets could detail which fees they are charging someone for a ticket, and accurately and adequately describe the choices available to a customer. That does not prevent the ticket vendor from overcharging the user. The underlying power imbalance results in a situation too unfair for interaction design to fix. If the ticket vendor is forced to use fair design, this shifts *some* power to the users – they will know they are being ripped off – but from there, further action will still be needed to address the underlying unfairness.

The realization that interaction design alone cannot solve all fairness problems is not surprising. A similar argument has been made in the past about standards proposed in FAccT papers. Concretely, the paper *"A Mulching Proposal"* satirically proposes an algorithm that would turn the elderly into animal food [45]. This algorithm would adhere to these standards proposed in FAccT papers. It would be non-discriminating and transparent, in sensible ways. None of this solves the underlying problem that the proposed algorithm would turn the elderly into animal food. This absurd example illustrates the limitations of narrow standards: turning the elderly into animal food is not a fair proposal, regardless of compliance with standards. This is similar for interaction design: if a situation is deeply unfair or unjust, it cannot be fixed with interaction design.

The fact that a fair pattern does not guarantee a fair design leads to the risk of *fairwashing*. For this concept, we borrow from greenwashing [22]. Greenwashing refers to a company communicating as though their practices are climate friendly, whereas their business practices are not. A related practice is pinkwashing, where businesses communicate as though they are friendly to the queer community, although they treat their queer employees poorly [6]. In the same manner, we can formulate fairwashing as a company communicating as though their practices are fair, whereas their business practices are not.

We can find an example that would seem to be fairwashing in the related field of fair algorithms. Van den Broek et al. [81] perform an ethnographic study of a multinational company deploying an AI system to assist in the hiring process. The application, which is not named in the article, claims to "eliminate bias and guarantee objective and role-specific evaluations". The company had deployed this application to work towards a fairer HR process. As Van den Broek et al. [81] describe in their article, the results were not without controversy. The HR team described that the application insufficiently catered to different contexts. Candidates figured out ways to game the system, thus introducing new difficulties in evaluation. Managers discovered that the algorithm rated a cherished intern terribly, leading the managers to challenge the validity of the algorithm. The developers of the AI claimed to have created a fair application, but in practice, many problems of traditional HR persisted even with use of the AI tool. If we try to establish a standard of fairness in design, we can imagine similar practices happening here.

It is difficult to evaluate design patterns as isolated concepts. If we inappropriately characterize design patterns as fair, companies might use this "fair" design pattern to justify their entire design as fair. We have established that fairness is contestable and contextual. As such, fairness is a difficult target to hit. We must be specific about context, with a clear notion of what fairness means in that context, before we can meaningfully call a design pattern fair.

4.2 Fair design alone cannot fix society's ills

Well-meaning developers could also run into problems by framing fairness too broadly. They could try to achieve fairness in the broadest sense, and only call a design fair if it achieves broad societal good: equity, or climate justice. An attempt to do societal good through interaction design might well be a nudge. A *nudge* is "any aspect of the choice architecture that alters people's behaviour in a predictable way, without forbidding any options or significantly changing their economic incentives" [78]. A fine is not a nudge, since it incurs a direct cost. Setting a default option is a nudge, since a default option does not change the possible options, but a default option makes one choice more accessible. Then, if we try to do societal good through changing the user interface, we are nudging.

Mani et al. [54] establish that the poor face more cognitive load than the rich. Steering design – even if this steering is done for good, i.e. nudging – can place extra cognitive burden on those affected [50, 58]. Steering design requires effort to circumvent, and as such, requires an 'effort tax' to avoid [75]. This means that steering design (even if intended for good) can harm the poor more than the rich. More zoomed out, if we design to achieve some external good, the cost of that will primarily be borne by those who cannot pay the

effort tax to avoid it. This means that those with time to spare will be able to reap the benefits of the measure, while being less affected by the cost [69].

Employing design for societal good runs the immediate risk of *responsibilisation*. Responsibilisation describes "the process whereby subjects are rendered individually responsible for a task which previously would have been the duty of another – usually a state agency – or would not have been recognized as a responsibility at all" [83]. If we focus on directing shoppers towards booking a more eco-friendly flight option, this seemingly puts responsibility for climate outcomes on the shopper. This is not entirely unreasonable; the individual does have responsibility for their own actions. However, climate change is a problem on a much broader scale than individual decision making. In order to tackle climate change, one requires policy, and collective action, rather than just individual choices. Fair Design could have a position in supplementing policy, but Fair Design should not be made the backbone of strategy against societal issues [67].

5 Where do we go from here?

Previously, we have found that the notion of fairness does not adequately address deceptive patterns. Fairness as moral desirability is a small, moving target, and trying to apply fairness to design patterns runs various risks. Yet, it is evidently possible to improve on current interface design. The concept of fairness as moral desirability seems to be a hindrance, more than a help, in providing practical guidance to interface designers. Fazelpour and Lipton [29] identify that practical guidance "requires understanding the relevant causal mechanisms that (i) account for present injustices; and (ii) govern the impact of proposed interventions" [29]. These criteria do not depend on whether this guidance meets some gold standard of fairness. Fazelpour and Lipton [29] argue for a non-ideal approach to improvement, using empirically informed, practical solutions as guidance, outside of an ideal standard of fairness. Just because a solution is not perfectly fair, does not mean that it is not worthwhile.

For example, the German Fair Consumer Contracts Act mandates that businesses that offer subscriptions to German consumers implement a "Two-Click Cancellation" interface [10, 27, 47]. More specifically, the act requires websites that offer subscriptions to include an easily accessible button labeled "terminate contracts here". This button must directly lead to a form where customers can cancel the subscription by filling in the required details and clicking a "terminate now" button. After the customer has clicked the button, the business must immediately confirm the cancellation.

This law provides a design pattern that makes canceling subscriptions easier compared to many practices used in the wild [71]. One could argue whether this design pattern meets some standard of fairness, or which considerations of fairness are appropriate here. Regardless of these questions, other lawmakers have followed Germany's lead: France has mandated three-click cancellation [24], and the European Union is introducing similar legislation [47].

It may also be possible to provide practical guidance by splitting fairness into its constituent parts. As described in Section 3.2, fairness is multifaceted, including values such as privacy and accessibility. Already, there are design patterns that enhance privacy, and design patterns that enhance accessibility. Armed with these design patterns as the building blocks for their designs, designers can directly address the specific harms caused by their systems, rather than trying to achieve a gold standard.

Lastly, in this work, we introduced the notion of *fairwashing*. Theoretically, it seems plausible that fairwashing can exist, but there does not appear to be a body of writing yet on whether users perceive a company as fairer when the company claims to be fair. We welcome more research on what users deem fair, and particularly, to what extent fairwashing could have an effect.

6 Conclusion

In this work, we have outlined potential difficulties in trying to develop fair patterns. We have demonstrated that fairness is a complex notion, as fairness depends on who is evaluating the design, and the context in which the design is deployed. This means that to call anything fair, one must detail under what conviction it is fair, and in which contexts it is fair. The moral value of a design pattern *in use* might be different from the moral value of the design pattern *in isolation*. This means that we cannot speak of fair patterns as such, but rather, of fair or unfair use of a given design pattern.

Additionally, we have illustrated when fair patterns cannot solve problems. Presenting unfair choices adequately does not make them fair. Fair patterns also cannot solve societal issues at large. In both cases, creating user interactions targeting improvement can be a worthwhile endeavor, but in support of policy targeting these issues — not standing alone.

As a closing note, we must not let the perfect become the enemy of the good. Even if we cannot create perfectly fair design patterns, imperfect incremental improvements are worthwhile to make user interaction better for all users.

Acknowledgments

We would like to thank the Lorentz Center for hosting the workshop that sparked this paper, all other workshop participants for the inspiring discussions, and Djoerd Hiemstra for his valuable input.

References

[1] Amurabi 2024. 2024. Fair Patterns - Design high growth products without legal risk. https://fairpatterns.com/
[2] Julio Abascal and Colette Nicolle. 2005. Moving towards inclusive design guidelines for socially and ethically aware HCI. *Interacting with computers* 17, 5 (2005), 484–505.
[3] C. Alexander, S. Ishikawa, and M. Silverstein. 1977. *A Pattern Language: Towns, Buildings, Construction.* OUP USA. https://books.google.nl/books?id=hwAHmktpk5IC
[4] Agathe Balayn, Christoph Lofi, and Geert-Jan Houben. 2021. Managing bias and unfairness in data for decision support: a survey of machine learning and data engineering approaches to identify and mitigate bias and unfairness within data management and analytics systems. *The VLDB Journal* 30, 5 (2021), 739–768.
[5] Dennis Benner, Sofia Marlena Schöbel, Andreas Janson, and Jan Marco Leimeister. 2022. How to achieve ethical persuasive design: A review and theoretical propositions for information systems. *AIS Transactions on Human-Computer Interaction* 14, 4 (2022), 548–577.
[6] Corinne E Blackmer. 2019. Pinkwashing. *Israel Studies* 24, 2 (2019), 171–181.
[7] Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. 2020. Language (technology) is power: A critical survey of" bias" in nlp. *arXiv preprint arXiv:2005.14050* (2020).
[8] Jeffrey Brand-Ballard. 2010. Treating Like Cases Alike. In *Limits of Legality: The Ethics of Lawless Judging.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195342291.003.0015 arXiv:https://academic.oup.com/book/0/chapter/276128565/chapter-ag-pdf/44535412/book_32868_section_276128565.ag.pdf

[9] Harry Brignull. 2010. Deceptive Design–formerly darkpatterns. org.

[10] Bundestag. 2021. Gesetz für Faire Verbraucherverträge. *Federal Gazette (Part I)* (2021), 3433–3435. Issue 53. full text publicly available (in German) at http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl121s3433.pdf

[11] 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 (Oct. 2016), 237–254. https://doi.org/10.1515/popets-2016-0038

[12] Ann Cavoukian. 2010. Privacy by design: the definitive workshop. A foreword by Ann Cavoukian, Ph. D. *Identity in the Information Society* 3, 2 (2010), 247–251.

[13] Lorena Sánchez Chamorro, Kerstin Bongard-Blanchy, and Vincent Koenig. 2023. Ethical Tensions in UX Design Practice: Exploring the Fine Line Between Persuasion and Manipulation in Online Interfaces. (2023).

[14] David Chaum. 1983. Blind signatures for untraceable payments. In *Advances in Cryptology: Proceedings of Crypto 82*. Springer, 199–203.

[15] David Chaum. 1985. Security without identification: Transaction systems to make big brother obsolete. *Commun. ACM* 28, 10 (1985), 1030–1044.

[16] David L Chaum. 1981. Untraceable electronic mail, return addresses, and digital pseudonyms. *Commun. ACM* 24, 2 (1981), 84–90.

[17] Richard J Chen, Judy J Wang, Drew FK Williamson, Tiffany Y Chen, Jana Lipkova, Ming Y Lu, Sharifa Sahai, and Faisal Mahmood. 2023. Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nature biomedical engineering* 7, 6 (2023), 719–742.

[18] Sünje Clausen, Julian Marx, Milad Mirbabaie, and Stefan Stieglitz. 2022. From Dark Patterns to Digital Sludging-Mapping the Ethical Debate on Controversial Persuasive System Design.. In *ICIS*.

[19] Damian Clifford and Jef Ausloos. 2018. Data protection and the role of fairness. *Yearbook of European Law* 37 (2018), 130–187.

[20] Michael Colesky, Jaap-Henk Hoepman, and Christiaan Hillen. 2016. A critical analysis of privacy design strategies. In *2016 IEEE security and privacy workshops (SPW)*. IEEE, 33–40.

[21] Gregory Conti and Edward Sobiesk. 2010. Malicious interface design: exploiting the user. In *Proceedings of the 19th international conference on World wide web*. 271–280.

[22] Sebastião Vieira de Freitas Netto, Marcos Felipe Falcão Sobral, Ana Regina Bezerra Ribeiro, and Gleibson Robert da Luz Soares. 2020. Concepts and forms of greenwashing: A systematic review. *Environmental Sciences Europe* 32 (2020), 1–12.

[23] 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*. 1–14.

[24] Direction de l'information légale et administrative (Premier ministre). 2024. Application de la résiliation « en 3 clics » : les dispositions spécifiques aux contrats d'assurance pouvant être conclus en ligne. https://www.service-public.fr/particuliers/actualites/A16455 Online; accessed 4-September-2024.

[25] Nick Doty and Mohit Gupta. 2013. Privacy design patterns and anti-patterns. In *Trustbusters Workshop at the Symposium on Usable Privacy and Security*.

[26] Bruno Dupont and Steven Malliet. 2021. Contextualizing Dark Patterns with the Ludeme Theory: A New Path for Digital Game Literacy? *Acta Ludologica* 4, 1 (2021), 4–22.

[27] Kristina Ehle and Stephen Kreß. 2021. "Two-Click Cancellation" Button – German Exceptionalism for Subscription Terminations. https://www.mofo.com/resources/insights/211006-new-two-click-cancellation-button#_ftn1 [Online; accessed 4-September-2024].

[28] Bart Engelen. 2019. Ethical criteria for health-promoting nudges: a case-by-case analysis. *The American journal of bioethics* 19, 5 (2019), 48–59.

[29] Sina Fazelpour and Zachary C Lipton. 2020. Algorithmic fairness from a non-ideal perspective. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. 57–63.

[30] Erich Gamma, Richard Helm, Ralph Johnson, and John Vlissides. 1995. *Design patterns: elements of reusable object-oriented software*. Pearson Deutschland GmbH.

[31] Damien Gerard. 2018. Fairness in EU competition policy: Significance and implications. , 211–212 pages.

[32] Inge Graef, Damian Clifford, and Peggy Valcke. 2018. Fairness and enforcement: bridging competition, data protection, and consumer law. *International Data Privacy Law* 8, 3 (2018), 200–223.

[33] 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*. 61–73.

[34] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018. The Dark (Patterns) Side of UX Design. In *CHI*. ACM, 534.

[35] Colin M. Gray, Cristiana Teixeira Santos, Nataliia Bielova, and Thomas Mildner. 2024. An Ontology of Dark Patterns Knowledge: Foundations, Definitions, and a Pathway for Shared Knowledge-Building. In *CHI*. ACM, 289:1–289:22.

[36] Paul Graßl, Hanna Schraffenberger, Frederik Zuiderveen Borgesius, and Moniek Buijzen. 2021. Dark and Bright Patterns in Cookie Consent Requests. *Journal of Digital Social Research* 3, 1 (Feb. 2021), 1–38. https://doi.org/10.33621/jdsr.v3i1.54

[37] Ben Green. 2019. "Good" isn't good enough. In *Proceedings of the AI for Social Good workshop at NeurIPS*, Vol. 17.

[38] Munawar Hafiz. 2006. A collection of privacy design patterns. In *Proceedings of the 2006 conference on Pattern languages of programs*. 1–13.

[39] R. Hes and John Borking. 1995. Privacy-Enhancing Technologies: The Path to Anonymity.

[40] Jaap-Henk Hoepman. 2014. Privacy design strategies. In *IFIP International Information Security Conference*. Springer, 446–459.

[41] Jaap-Henk Hoepman. 2024. Do fair design patterns exist? [Blog post]. https://blog.xot.nl/2024/02/07/do-fair-design-patterns-exist/index.html

[42] Kenneth Hullett and Jim Whitehead. 2010. Design patterns in FPS levels. In *Proceedings of the Fifth International Conference on the Foundations of Digital Games* (Monterey, California) *(FDG '10)*. Association for Computing Machinery, New York, NY, USA, 78–85. https://doi.org/10.1145/1822348.1822359

[43] Peter Hustinx. 2010. Privacy by design: delivering the promises. *Identity in the Information Society* 3, 2 (Aug. 2010), 253–255. https://doi.org/10.1007/s12394-010-0061-z

[44] Duha Ibdah, Nada Lachtar, Satya Meenakshi Raparthi, and Anys Bacha. 2021. "Why Should I Read the Privacy Policy, I Just Need the Service": A Study on Attitudes and Perceptions Toward Privacy Policies. *IEEE Access* 9 (2021), 166465–166487. https://doi.org/10.1109/ACCESS.2021.3130086

[45] Os Keyes, Jevan Hutson, and Meredith Durbin. 2019. A mulching proposal: Analysing and improving an algorithmic system for turning the elderly into high-nutrient slurry. In *Extended abstracts of the 2019 CHI conference on human factors in computing systems*. 1–11.

[46] Th Kirat, Olivia Tambou, Virginie Do, and Alexis Tsoukiàs. 2023. Fairness and Explainability in Automatic Decision-Making Systems. A challenge for computer science and law. *EURO journal on decision processes* 11 (2023), 100036.

[47] Nathalie Koch and Richard Gläser. 2024. Germany: Two years of cancellation buttons – status quo and new developments. https://www.taylorwessing.com/en/insights-and-events/insights/2024/06/germany-two-years-of-cancellation-buttons Online; accessed 4-September-2024.

[48] Alina Köchling and Marius Claus Wehner. 2020. Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research* 13, 3 (2020), 795–848.

[49] Christopher Kuner, Lee A Bygrave, Christopher Docksey, Laura Drechsler, and Luca Tosoni. 2021. The EU general data protection regulation: A commentary/update of selected articles. *Update of Selected Articles (May 4, 2021)* (2021).

[50] Matthias Lehner, Oksana Mont, and Eva Heiskanen. 2016. Nudging–A promising tool for sustainable consumption behaviour? *Journal of cleaner production* 134 (2016), 166–177.

[51] Luanna Lopes Lobato, Eduardo B Fernandez, and Sérgio Donizetti Zorzo. 2009. Patterns to support the development of privacy policies. In *2009 International Conference on Availability, Reliability and Security*. IEEE, 744–749.

[52] Jamie Luguri and Lior Jacob Strahilevitz. 2021. Shining a light on dark patterns. *Journal of Legal Analysis* 13, 1 (2021), 43–109.

[53] Gianclaudio Malgieri. 2020. The concept of fairness in the GDPR: a linguistic and contextual interpretation. In *FAT* '20: Conference on Fairness, Accountability, and Transparency, Barcelona, Spain, January 27-30, 2020*, Mireille Hildebrandt, Carlos Castillo, L. Elisa Celis, Salvatore Ruggieri, Linnet Taylor, and Gabriela Zanfir-Fortuna (Eds.). ACM, 154–166. https://doi.org/10.1145/3351095.3372868

[54] Anandi Mani, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. 2013. Poverty impedes cognitive function. *science* 341, 6149 (2013), 976–980.

[55] Floyd Marinescu. 2002. *EJB design patterns*. Wiley New York.

[56] Andrei Marmor. 2005. SHOULD LIKE CASES BE TREATED ALIKE? *Legal Theory* 11, 1 (2005), 27–38. https://doi.org/10.1017/S1352325205050020

[57] Arunesh Mathur, Gunes Acar, Michael J Friedman, Eli Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan. 2019. Dark patterns at scale: Findings from a crawl of 11K shopping websites. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–32.

[58] 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*. 1–18.

[59] Christian Meske and Ireti Amojo. 2020. Ethical guidelines for the construction of digital nudges. *arXiv preprint arXiv:2003.05249* (2020).

[60] Thomas Mildner, Merle Freye, Gian-Luca Savino, Philip R Doyle, Benjamin R Cowan, and Rainer Malaka. 2023. Defending against the dark arts: recognising dark patterns in social media. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 2362–2374.

[61] Jakob Nielsen. 1994. 10 Usability Heuristics for User Interface Design. https://www.nngroup.com/articles/ten-usability-heuristics/

[62] Donald A Norman and Stephen W Draper. 1986. *User centered system design; new perspectives on human-computer interaction*. L. Erlbaum Associates Inc.

[63] Marie Potel-Saville and Mathilde Da Rocha. 2023. From Dark Patterns to Fair Patterns? Usable Taxonomy to Contribute Solving the Issue with Countermeasures. (2023).

[64] Karen Renaud and Verena Zimmermann. 2018. Ethical guidelines for nudging in information security & privacy. *International Journal of Human-Computer Studies* 120 (2018), 22–35.

[65] Hauke Sandhaus. 2023. *Brightpatterns.org: A Front Page to Define the Term and Collect Examples.* LIC, New York City. https://www.brightpatterns.org Archived at https://perma.cc/K8MB-47TR on February 7, 2024.

[66] Hauke Sandhaus. 2023. Promoting Bright Patterns. CHI '23 Workshop: Designing Technology and Policy Simultaneously. arXiv:2304.01157 [cs.HC]

[67] Andreas T Schmidt and Bart Engelen. 2020. The ethics of nudging: An overview. *Philosophy compass* 15, 4 (2020), e12658.

[68] Hanna Schraffenberger, Raphael Gellert, Colin M. Gray, Arianna Rossi, and Cristiana Santos. 2024. Fair patterns for online interfaces [Workshop]. https://www.lorentzcenter.nl/fair-patterns-for-online-interfaces.html Accessed: 26 July 2024.

[69] Christian Schubert. 2017. Green nudges: Do they work? Are they ethical? *Ecological Economics* 132 (Feb. 2017), 329–342. https://doi.org/10.1016/j.ecolecon.2016.11.009

[70] Andrew D Selbst, Danah Boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*. 59–68.

[71] Ashley Sheil, Gunes Acar, Hanna Schraffenberger, Raphael Gellert, and David Malone. 2024. Staying at the Roach Motel: Cross-Country Analysis of Manipulative Subscription and Cancellation Flows. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 298, 24 pages. https://doi.org/10.1145/3613904.3642881

[72] Marie-Sophie Simon, Hanna Schraffenberger, and Raphaël Gellert. 2024. Influencing User Decisions: Dilemmas in Designing Online Interfaces. In *Adjunct Proceedings of the 2024 Nordic Conference on Human-Computer Interaction (NordiCHI Adjunct 2024)* (Uppsala, Sweden). Association for Computing Machinery, New York, NY, USA, 6 pages. https://doi.org/10.1145/3677045.3685441

[73] Nathalie A Smuha. 2019. The EU approach to ethics guidelines for trustworthy artificial intelligence. *Computer Law Review International* 20, 4 (2019), 97–106.

[74] Daniel J Solove. 2005. A taxonomy of privacy. *U. Pa. l. Rev.* 154 (2005), 477.

[75] Cass R Sunstein and Lucia A Reisch. 2013. Green by Default. *Kyklos* 66, 3 (2013).

[76] David Tannenbaum, Craig R Fox, and Todd Rogers. 2017. On the misplaced politics of behavioural policy interventions. *Nature Human Behaviour* 1, 7 (2017), 0130.

[77] A. Terpstra, P. Graßl, and H. Schraffenberger. 2021. Think before you click: how reflective patterns contribute to privacy [Position paper]. In *Workshop "What Can CHI Do About Dark Patterns?" at the CHI Conference on Human Factors in Computing Systems (CHI'21)*. https://repository.ubn.ru.nl/handle/2066/246490

[78] Richard H Thaler and Cass R Sunstein. 2021. *Nudge: The final edition*. Yale University Press.

[79] Jenifer Tidwell. 2005. *Designing interfaces: Patterns for effective interaction design*. " O'Reilly Media, Inc.".

[80] Jelmer Tiemessen, Hanna Schraffenberger, and Gunes Acar. 2023. The Time is Ticking: The Effect of Limited Time Discounts on Consumers' Buying Behavior and Experience. In *CHI Extended Abstracts*. ACM, 277:1–277:11.

[81] Elmira Van den Broek, Anastasia Sergeeva, and Marleen Huysman. 2020. Hiring algorithms: An ethnography of fairness in practice. In *40th international conference on information systems, ICIS 2019*. Association for Information Systems, 1–9.

[82] Sandra Wachter, Brent Mittelstadt, and Chris Russell. 2021. Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review* 41 (2021), 105567.

[83] Alison Wakefield and Jenny Fleming. 2008. *The Sage dictionary of policing*. SAGE Publications Ltd.

[84] Yifan Wang, Weizhi Ma, Min Zhang, Yiqun Liu, and Shaoping Ma. 2023. A survey on the fairness of recommender systems. *ACM Transactions on Information Systems* 41, 3 (2023), 1–43.

[85] Hilde Weerts, Raphaële Xenidis, Fabien Tarissan, Henrik Palmer Olsen, and Mykola Pechenizkiy. 2023. Algorithmic unfairness through the lens of EU non-discrimination law: Or why the law is not a decision tree. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 805–816.

[86] Chris Willett. 2010. Fairness and consumer decision making under the unfair commercial practices directive. *Journal of Consumer Policy* 33 (2010), 247–273.

[87] José P Zagal, Staffan Björk, and Chris Lewis. 2013. Dark patterns in the design of games. In *Foundations of Digital Games 2013*.