

Layered Analysis of Persuasive Designs: A Framework for Identification and Autonomy Evaluation of Dark Patterns

Sanju Ahuja^{1,*} and Jyoti Kumar¹

¹ Indian Institute of Technology Delhi, New Delhi 110016, India

Abstract

The use of persuasive designs to influence user behavior has become ubiquitous in digital contexts. This has led to the emergence of ethically dubious design practices, highlighted in literature under the paradigm of 'dark patterns'. Various taxonomies of dark patterns have been reported, however, there is a lack of frameworks which capture how such designs can be embedded not just within user interfaces, but also in functionality and strategy of digital systems. This paper proposes a framework for a Layered Analysis of Persuasive Designs, grounded in Garrett's five layered model of user experience (UX) design and Fogg's Behavior Design Model. The framework identifies a toolkit of 48 design elements which can be used to operationalize problematic persuasion in digital contexts. It also highlights the autonomy impact of each design element. The proposed framework can help designers and policymakers identify (potential) dark patterns within digital systems and evaluate them from an autonomy perspective.

Keywords

dark patterns, persuasive design, user experience design, user autonomy, ethics

1. Introduction

Persuasive design refers to design that intends to influence users' decision making and behavior [1-3]. The use of persuasion has become ubiquitous in digital contexts [4-6]. Persuasion can be designed with varied intent. For example, mobile applications can motivate users towards health and fitness goals and digital well-being features can help users manage their screen time [7,8]. It can also be designed purely for commercial gain, such as to increase revenue, gather user information and maximize user engagement [9]. This has led to the emergence of ethically dubious design practices, studied in literature under the paradigm of 'dark patterns'. These designs can trick, coerce or manipulate users into acting in ways not intended by them or not in their best interests [10,11].

The paradigm of dark patterns has captured the phenomenon of problematic persuasive designs within commercial digital systems, such as e-commerce, mobile games, social media, online privacy, online advertising and home robots [12-17]. The act of persuasion, even with an intent of beneficence, is often in tension with the values of autonomy [18-20].

Mobilizing Research and Regulatory Action on Dark Patterns and Deceptive Design Practices Workshop at CHI conference on Human Factors in Computing Systems, May 12, 2024, Honolulu, HI (Hybrid Workshop)

* Corresponding author.

✉ sanju.ahuja@design.iitd.ac.in (S. Ahuja); jyoti@design.iitd.ac.in (J. Kumar)

🆔 0000-0002-5178-5077 (S. Ahuja); 0000-0002-3810-3262 (J. Kumar)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Hence, concerns with dark patterns frequently pertain to user autonomy [11]. In addition, they can be accompanied by a loss of user welfare, such as unexpected financial loss, invasion of privacy, cognitive burden, addiction to social media and games, loss of health, time and confidence [11,21,22]. They can also cause structural harms, undermine competition, prevent price comparison and reduce trust in the market [11,23-27].

While important advances have been made to taxonomize dark patterns in various contexts, the literature lacks frameworks which can help designers and policymakers identify such designs within any digital system and evaluate them from an autonomy perspective. To address this research gap, we propose a framework for a Layered Analysis of Persuasive Designs. In this framework, we take the approach of understanding how dark designs are operationalized through ‘interaction design elements’. We identify 48 potentially problematic persuasive design elements and situate them within existing theories and models of user experience (UX) design and behavior change. We also highlight the autonomy impact of each design element. We argue that this framework can help designers and policymakers (a) identify problematic persuasive designs in digital contexts, and (b) evaluate these designs from an autonomy perspective.

2. Background

2.1. Dark Patterns

The term dark patterns, coined by UX practitioner Harry Brignull [13], was originally defined as “*tricks used in websites and apps that make you do things that you didn't mean to, like buying or signing up for something.*” Since then, literature has developed upon the concept with rich descriptive contributions, identifying and taxonomizing dark patterns in various contexts, such as online privacy [12], e-commerce [15] and mobile games [17]. There have also been attempts to consolidate this literature at a more abstract level. Mathur et al. [11] systematically reviewed dark patterns literature and categorized them into two strategic approaches: *Modifying the decision space* and *Manipulating the information flow*, with six subcategories. Ahuja & Kumar [10] conducted a systematic review of dark patterns and categorized them into 25 dark strategies. They further segregated the dark strategies into seven ethical concerns grounded in four aspects of autonomy. Gray et al. [28] conducted a systematic review of dark patterns and segregated them into high-level, meso-level and low-level patterns, based on the level of abstraction in the definition of each pattern.

2.2. Autonomy Concerns

Discourses on the ethics of persuasion have widely articulated concerns about autonomy. Berdichevsky and Neuenschwander [29] were perhaps the first to argue that persuasive digital technologies should respect individual autonomy and should be used to promote ethical outcomes. Fogg [1] raised concerns about particular strategies such as using emotions, deception, coercion, operant conditioning and surveillance. Verbeek [6] argued that ambient intelligence can create radically new persuasions and may encroach upon users’ everyday activities and choice processes. Nagenborg [30] discussed freedom concerns with persuasion based on surveillance, even when privacy concerns are minimal.

Susser [31] discussed how behavioral advertising can be used for online manipulation and articulated its harm to individual autonomy. Latham [32], from the UX Collective, connected ethical design to personal freedom and argued that subtle manipulations in advertisements and digital media may condition users' choices. Concerns about autonomy have also been made explicit in the dark patterns literature. Brignull [13] defined dark patterns as tricks that 'make you do things you didn't mean to' and Bösch et al. [12] argued that dark patterns trick users into performing 'unintended' and 'unwanted' actions. Others have argued that dark patterns deceive, manipulate and mislead users [14,16,33].

Autonomy is understood in philosophical literature as an individual's right to self-governance [34]. It is constituted by the freedom to be the person one wants to be and to pursue one's goals without unjustifiable hindrances or interferences [35]. The value of autonomy subsumes different concepts or notions. In behavioral economics, a systematic review of literature showed that three conceptualizations of autonomy underlie the ethical debate about nudging: agency, freedom of choice and self-constitution (or authenticity) [36]. A systematic review of dark patterns literature showed that four conceptualizations of autonomy underlie the ethical concerns about dark patterns: agency, freedom of choice, control and independence [10]. Together, the literature points to five autonomy related ethical concerns: *Agency*, *Freedom of choice*, *Control*, *Independence* and *Authenticity* [37].

Agency refers to an individual's capacity to choose and decide [36,38]. It involves the ability to reason about and reflect upon one's preferences and choices. Manipulation and deception can pose threats to agency, such that an individual's decision making and reasoning capacities are bypassed by external influences. **Freedom of choice** includes the availability of options and an absence of undue restrictions [30]. It means that individuals have a practical, and not just a theoretical availability of choices [39]. Freedom can be undermined through force, pressure, coercion [40], and by making access to relevant choices impossible or costly [41]. **Control** involves that individuals are provided with the opportunity to make decisions on their own behalf or consent to decisions that have relevance for them [10,37]. Threats to control manifest in the form of non-consensual decision making, such as taking an action on behalf of the individual. **Independence** is grounded in the notion of non-reliance [10,37]. Lack of independence implies a lack of self-regulation of behavior. This can manifest as obsessions, addictions, dependencies and compulsions [22]. Threats to independence are external influences which act on an individual's limited propensity for self-control and foster behaviors that the individual is unable to self-regulate effectively. **Authenticity** includes concerns dealing with a person's identity [36]. It requires an individual to act for reasons that can be considered their own [42,43]. Hence, it is undermined when individuals are made to endorse desires and values or act in ways that cannot truly be attributed to their own self.

2.3. Models of Interaction Design and UX Design

Various models of interaction design and user experience (UX) design have been proposed in literature. These models can help understand the different elements of any interactive system as well as the elements of user experience with interactive systems. This section highlights some prevalent models from literature.

Norman [44] conceptualized the interaction between a user and a system as an execution-evaluation cycle. The model argues that while interacting with a system, the user first sets a goal and executes it on the system (execution component), and then evaluates the outcomes of the system's actions (evaluation component). Abowd and Beale [45] extended the model to make the role of user interface (UI) explicit within this interaction cycle. In the book *Designing Interactions* [46], Gillian Crampton Smith proposed that there are four dimensions to user experience design within interactive systems. A fifth dimension was added by Silver [47]. These are: Words (1D), Visual Representations (2D), Physical Objects / Space (3D), Time (4D) and Behavior (5D). Morville [48] conceptualized user experience as a 7-Facet Honeycomb. He argued that a system's experience depends on whether it is useful (offers a function or a purpose), usable (effective and efficient to use), findable (easy to locate content and products), credible (trustworthy), accessible (allows access for users of all capabilities), desirable (through branding, visuals, and aesthetics which evoke appreciation), and valuable (overall commercial and customer value, which is determined by the other six factors).

Garrett [49] proposed a model to develop the user experience of a digital product in layers. The model proposed five layers of UX design: Surface, Skeleton, Structure, Scope and Strategy. The surface is the most concrete layer of a product comprising of its visual design. However, the visual design of a product is a culmination of product decisions in the layers underneath. The UX of a product can be understood as progressing from the abstract 'strategy' to the concrete 'surface', with intermediate layers in between. **Surface** comprises of visible elements like colors, images and typography. Beneath the surface is the **Skeleton** of the product, which accounts for the placement of visible elements such as buttons, tabs, images and blocks of text. This skeleton is an expression of the **Structure** of a product, which consists of its interaction design and information architecture. The **Scope** defines the functions and features of a product. The scope is fundamentally determined by **Strategy**, which comprises of user needs and product objectives. Product objectives are usually determined by business goals.

This paper argues that problematic persuasion can emerge at different layers of UX design. It is not only implemented using visible elements such as colors and sizes of buttons, but also through invisible elements such as trust and social influence. Problematic persuasion can be a part of the design strategy itself, such as for products which are solely designed to maximize engagement. Similar arguments have been articulated previously. Fogg [1] argued that interactive products can be designed with an overall persuasive intent, labelling this approach as macrosuasion. Westin and Chiasson [16] referred to the systematic use of designs to serve a problematic intent, labelling the issue as dark infrastructures. Garrett's layered model [49] is one of the few interaction design models that explicitly capture these phenomena. It highlights the role of visible and invisible design elements in the user experience of a digital system. Hence, the paper grounds its proposed framework within Garrett's five layers of UX design.

2.4. Models of Persuasive Design

Several models have been proposed in literature to help designers integrate persuasion within interactive systems. These include the Behavior Design Model [2], Behavior Change

Wheel [50], Persuasive Systems Design Model [3], Influence Grid [51], Digital Nudging Process Model [52], and Taxonomy of Choice Architecture Interventions [53,54]. Several of these take the form of process models or suggest toolkits for the design of behavior change interventions. The Behavior Design Model [2] is one of the few conceptual models which highlight the psychological constructs of behavior change. It has been widely adopted in literature to guide the development of behavior change interventions [55,56]. Hence, the paper grounds its proposed framework within the Behavior Design Model. The model argues that for a user to perform a desired behavior, they must be (a) sufficiently motivated, (b) have an ability, and (c) be triggered. Hence, behavior change can be brought about by affecting one of three factors: **Motivation, Ability and Triggers**.

This model posits that a designer can increase the likelihood of a user performing a behavior by enhancing their motivation. Similarly, a behavior can be prevented by reducing motivation. Three core motivators with two opposing dimensions were discussed in the model: pleasure / pain, hope / fear, and acceptance / rejection. A designer can also affect behavior change by enhancing or reducing users' ability to perform a behavior. To produce a target behavior, it can be simplified in terms of time, money, physical effort, mental effort and routineness. Similarly, to prevent a target behavior, it can be complicated on the same factors. Lastly, a designer can trigger users to perform the desired behavior. Triggers go by various names such as prompts, cues, calls to action, etc. Triggers provide an opportunity for the user to perform a behavior. Triggers can be neutral and can act as mere reminders or signals. However, they can be designed in tandem with a motivational influence (sparks) or an ability influence (facilitators) to maximize their effectiveness. On the other hand, they can also cause users to act on impulse or be perceived as a nuisance.

3. Layered Analysis of Persuasive Designs

This section proposes a framework for a Layered Analysis of Persuasive Designs. The following steps were taken to develop this framework. These steps are depicted in Figure 1.

1. **Corpus generation of design elements:** The first step was to generate a corpus of design elements that can be used for persuasive design in digital contexts. A preliminary corpus of design elements was generated in a brainstorming session by the first author with a design professional (male, 28 years). This design professional had 7+ years of industry experience and was currently employed as a senior product manager in a software company. Different types of design elements were generated, such as visual elements on the screen (text font, color of buttons, etc.), elements pertaining to navigation (sequence of information, timing of information, number of steps, etc.), and elements of functionality (choices, prompts, defaults, etc.).
2. **Derivation of design elements which impact autonomy:** After this initial corpus was generated, the first author devised the possibilities of how the dark strategies reported in [10] could be implemented using this corpus of design elements. For example, 'small font sizes' (design element) can be used for 'information hiding' (dark strategy), which can lead to 'inadequate information' (ethical concern). This can impact 'agency' (aspect of autonomy). This step led to a corpus of potentially

problematic design elements such as ‘small font sizes which hide information’, ‘multiple navigation steps which hide information’, ‘preselected or default choices which bypass consent’, etc., along with their underlying autonomy related ethical concerns. In this step, problematic design elements pertaining to the *Authenticity* conceptualization also emerged, which had not been reported in [10].

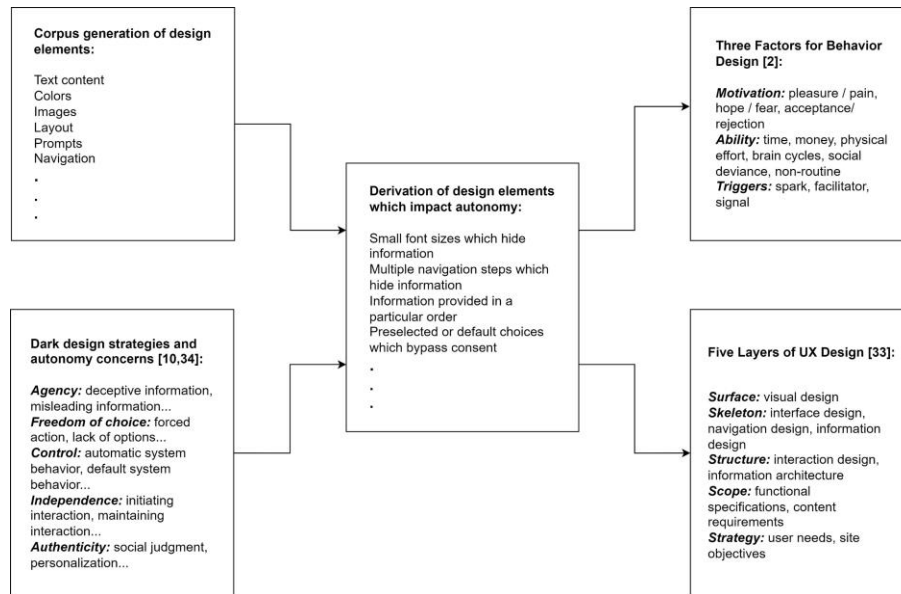


Figure 1: Research Methodology for Framework Development

3. **Pilot testing:** Pilot testing was done with the help of five PhD researchers in the field of user experience (UX) design. The author introduced the researchers to the concept of persuasive design and dark patterns. The researchers assessed different digital platforms and identified design features which influenced users to spend money, time or share personal information on the platforms. They were provided with the corpus of potentially problematic design elements generated in the previous step as a guide. Feedback was collected to understand if any design elements were missing in the corpus and if the descriptions of the elements were clear and understandable. Based on this feedback, the corpus was iterated. Some design elements were combined, some were revised and new elements were added. This led to a corpus of 49 design elements.
4. **Coding and expert verification:** Each design element was coded based on (a) which layer of UX design the persuasion was embedded in: Surface, Skeleton, Structure, Scope, and Strategy [49], and (b) which behavioral factor it affected: Motivation, Ability, and Triggers [2]. The coding was done by the first author and verified by an additional expert (female, 39 years) with a PhD in user experience design. The expert was versed with the five layers of UX design and was introduced to the Behavior Design Model by the author. To verify the codes, coding guides were provided for both models, reported in the Appendix. Any initial disagreements in the codes were resolved through mutual discussion. This expert also gave feedback on

the descriptions of the design elements and helped eliminate one redundant element. Hence, a final corpus of 48 design elements was included in the framework.

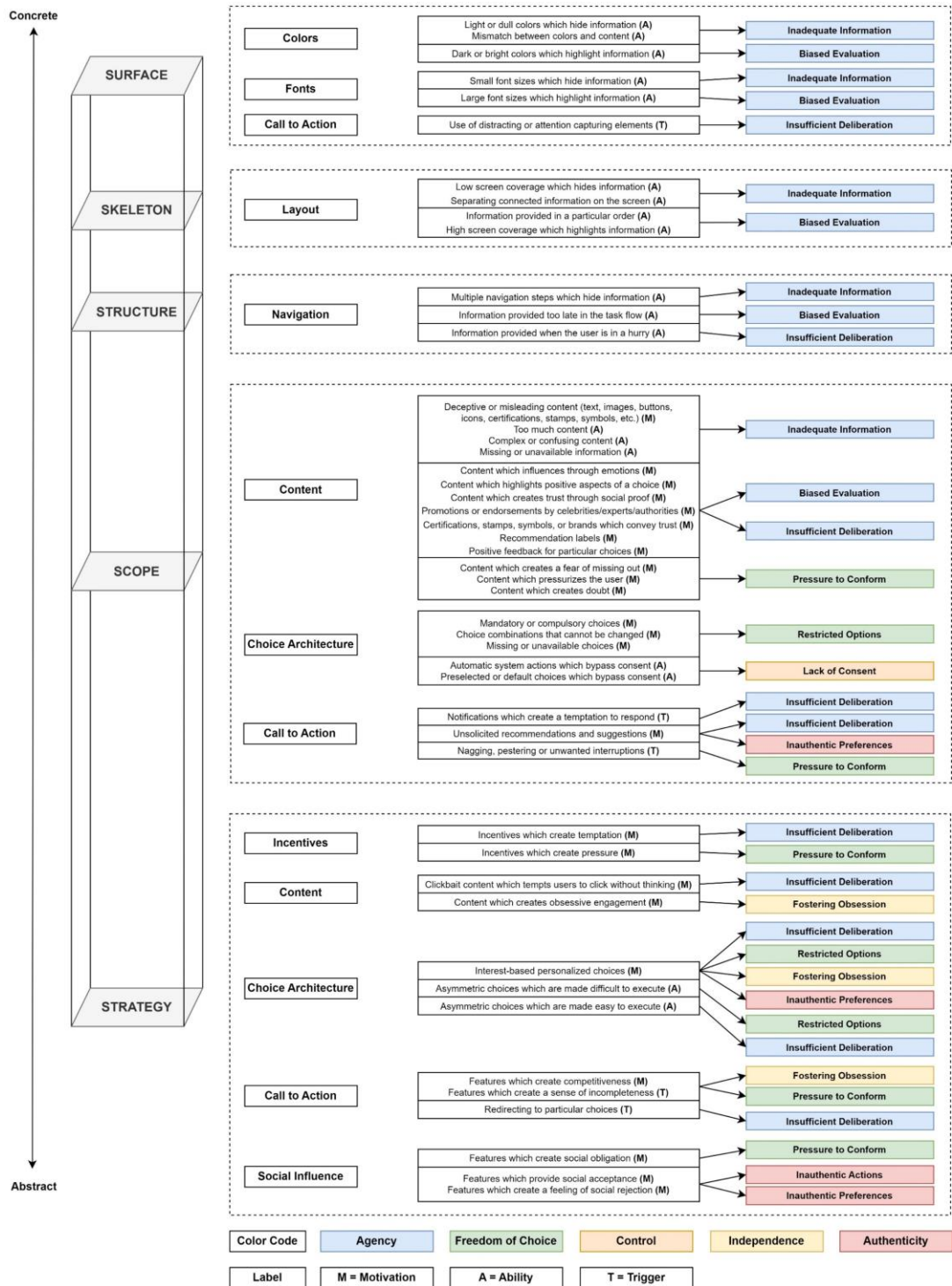


Figure 2: Proposed Framework for Layered Analysis of Persuasive Designs

5. **Proposed framework:** A visual illustration of the framework was created (Figure 2). The 48 persuasive design elements and the autonomy concerns associated with each element are reported in the framework. These elements are structured within the five layers of UX design, beginning from the Surface layer (for example, *'light or dull colors which hide information'*) and culminating in the Strategy layer (for example, *'features which create a feeling of social rejection'*). Whether these elements are designed to affect Motivation, Ability or Triggers is also reported in the figure (for example, *'light or dull colors which hide information'* leverage ability, and *'features which create a feeling of social rejection'* leverage motivation).

4. Discussion

The proposed framework has been developed with the intent of helping technology designers and policymakers identify persuasive designs in digital contexts and evaluate their autonomy impact. This section highlights the implications of this framework.

First, by contextualizing problematic design practices within five layers of UX design, the framework highlights how deep persuasion is embedded into a digital system. Persuasive designs in shallow layers (Surface / Skeleton / Structure) typically impact singular instances of decision making (also known as microsuasion [1]). They are also constituted by visible user interface elements. However, persuasion in deeper layers (Scope / Strategy) may not appear problematic without taking the invisible elements of product functionality or strategy into account. Second, the framework highlights issues pertaining to five aspects of autonomy. In previous literature, dark patterns which impact agency, freedom of choice and control have been frequently discussed [10]. Independence related issues have received less attention [10]. Authenticity related issues have not been discussed in dark patterns literature, even though they prominently appear in other contexts [36,57,58]. This is potentially because independence and authenticity are often impacted by relatively abstract or invisible aspects of design (also understood as macrosuasion by [1] and dark infrastructures by [16]). For example, social media can impact authenticity by creating a need for social validation and fear of social rejection through likes, comments and shares. However, the features themselves do not appear problematic unless the strategy is taken into account. Hence, the framework makes explicit autonomy concerns embedded in deeper layers of UX design. Third, by contextualizing problematic design practices within the Behavior Design Model [2], the framework highlights autonomy considerations in the persuasive design process. Hence, when practitioners use this model to design interventions, this framework can help them evaluate the autonomy impact of their designs.

The proposed framework includes a toolkit of 48 problematic design elements, largely from websites and mobile applications, which are a form of visual design. During the course of framework development, iterative pilot testing was done to ensure that the list of 48 design elements was as exhaustive as possible. Despite these efforts, it is not claimed that this is an exhaustive corpus. Newer design elements can be added to this corpus while preserving the underlying theoretical models that the framework is based on. In addition, the framework can be expanded further to include design elements from non-visual modalities of interaction. It is important to note that each of the design elements included

in the framework may be problematic to a different degree. User studies indicate that some of these design elements may be inherently more problematic than others [33], for example, *'Light or dull colors which hide information'*, *'Deceptive or misleading content'*, and *'Nagging, pestering or unwanted interruptions'*. However, the ethics of many design elements can also depend upon the context. For example, *'Mandatory or compulsory choices'* may be more problematic if the restrictions are unfair, *'Information provided in a particular order'* may be problematic if the design leads to poor outcomes, and so on. In addition, users' tolerance towards persuasive designs may also depend on culture. For example, Fogg [1] argued that Westerners might value individual freedoms while Easterners might prioritize institutional efficiency and collective power. Hence, the aim of this framework is to help designers identify persuasive designs in the digital context which have the potential to impact user autonomy. The framework does not aid in weighing ethical and practical tradeoffs to make conclusive normative judgments.

5. Conclusions and Future Work

This paper has proposed a framework for the normative assessment of persuasive designs from an autonomy perspective. The framework makes explicit the link between theoretical autonomy concerns and the applied form of interaction design elements (such as Colors, Fonts, Call to Action, Layout, Navigation, Content, Choice Architecture, Incentives and Social Influence). Hence, in a normative assessment of any digital system, it can guide designers to identify (potential) dark patterns. It can also be useful in HCI training and education, where it can supplement theoretical perspectives on ethics in a design classroom, by highlighting how different design elements create ethical concerns. This approach can ensure that students have both a theoretical and a pragmatic base of ethics knowledge which is grounded in fundamental principles but also aided by a toolbox of problematic design practices. Lastly, this framework can be useful to policymakers, who may not have formal training in UX design, behavior design or ethics. The toolbox of 48 design elements can be used even without a knowledge of the theoretical models underpinning the framework.

In future studies, we plan to draw comparisons between training designers with the proposed framework and training them with existing approaches (such as a theoretical knowledge of dark patterns and taxonomies of dark patterns). These studies can demonstrate the effectiveness of the framework for (a) helping designers identify persuasive design elements in digital contexts, and (b) for discerning the nature of the autonomy impact of individual design elements. This framework also has the potential to be developed into an evaluation tool. It consolidates problematic persuasive designs into five high-level autonomy conceptualizations. Hence, the framework can help in the development of theoretically grounded criteria to measure the degree of the impact of persuasive digital systems on different aspects of autonomy.

Acknowledgements

This research was supported by the Prime Minister's Research Fellowship granted by the Ministry of Education, Government of India.

References

- [1] B. J. Fogg, *Persuasive technology: using computers to change what we think and do*, Morgan Kaufmann Publishers, San Francisco, CA, USA, 2002.
- [2] B. J. Fogg, A behavior model for persuasive design, in: *Proceedings of the 4th International Conference on Persuasive Technology, Persuasive '09*, Association for Computing Machinery, New York, NY, USA, 2009, Article 40, pp. 1–7. doi:10.1145/1541948.1541999
- [3] H. Oinas-Kukkonen, M. Harjumaa, Persuasive systems design: Key issues, process model, and system features, *Communications of the Association for Information Systems* 24 (2009) 485–500. doi:10.17705/1CAIS.02428
- [4] I. Adaji, M. Adisa, A Review of the Use of Persuasive Technologies to Influence Sustainable Behaviour, in: *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization, UMAP '22 Adjunct*, Association for Computing Machinery, New York, NY, USA, 2022, pp. 317–325. doi:10.1145/3511047.3537653
- [5] S. Berkovsky, J. Freyne, H. Oinas-Kukkonen, Influencing Individually: Fusing Personalization and Persuasion, *ACM Transactions on Interactive Intelligent Systems* 2 (2012). doi:10.1145/2209310.2209312
- [6] P. P. Verbeek, Ambient Intelligence and Persuasive Technology: The Blurring Boundaries Between Human and Technology, *Nanoethics* 3 (2009) 231–242. doi:10.1007/s11569-009-0077-8
- [7] J. Matthews, K. T. Win, H. Oinas-Kukkonen, M. Freeman, Persuasive technology in mobile applications promoting physical activity: a systematic review, *Journal of Medical Systems* 40 (2016) 1–13. doi:10.1007/s10916-015-0425-x
- [8] L. S. Sullivan, P. Reiner, Digital Wellness and Persuasive Technologies, *Philosophy & Technology* 34 (2021) 413–424. doi:10.1007/s13347-019-00376-5
- [9] C. M. Gray, Y. Kou, B. Battles, J. Hoggatt, A. L. Toombs, The Dark (Patterns) Side of UX Design, in: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI '18*, Association for Computing Machinery, New York, NY, USA, 2018, Paper 534, pp. 1–14. doi:10.1145/3173574.3174108
- [10] S. Ahuja, J. Kumar, Conceptualizations of user autonomy within the normative evaluation of dark patterns, *Ethics and Information Technology* 24 (2022). doi:10.1007/s10676-022-09672-9
- [11] A. Mathur, M. Kshirsagar, J. Mayer, What Makes a Dark Pattern... Dark? Design Attributes, Normative Considerations, and Measurement Methods, in: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, CHI '21*, Association for Computing Machinery, New York, NY, USA, 2021, Article 360, pp. 1–18. doi:10.1145/3411764.3445610
- [12] C. Bösch, B. Erb, F. Kargl, H. Kopp, S. Pfattheicher, Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns, *Proceedings on Privacy Enhancing Technologies* 4 (2016) 237–254. doi:10.1515/popets-2016-0038
- [13] H. Brignull, What are dark patterns?, 2010. URL: <https://www.darkpatterns.org/>

- [14] C. Lacey, C. Caudwell, Cuteness as a 'Dark Pattern' in Home Robots, in: Proceedings of 14th ACM/IEEE International Conference on Human-Robot Interaction, HRI '19, Daegu, South Korea, 2019, pp. 374-381. doi:10.1109/HRI.2019.8673274
- [15] A. Mathur, G. Acar, M. J. Friedman, E. Lucherini, J. Mayer, M. Chetty, A. Narayanan, Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites, Proceedings of the ACM on Human-Computer Interaction 3 (2019). doi:10.1145/3359183
- [16] F. Westin, S. Chiasson, "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, CHI '21, Association for Computing Machinery, New York, NY, USA, 2021, Article 550, pp. 1-15. doi:10.1145/3411764.3445104
- [17] J. P. Zagal, S. Björk, C. Lewis, Dark patterns in the design of games, in: Proceedings of Foundations of Digital Games, FDG '13, Society for the Advancement of the Science of Digital Games, Santa Cruz, CA, 2013.
- [18] A. Maclean, Autonomy, Consent and Persuasion, European Journal of Health Law 13 (2006) 321-338. doi:10.1163/157180906779160274
- [19] N. Eyal, Informed Consent, in: E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy, Spring 2019 Edition, 2019. URL: <https://plato.stanford.edu/archives/spr2019/entries/informed-consent/>
- [20] D. A. Strauss, Persuasion, Autonomy, and Freedom of Expression, in: L. J. Alexander (Ed.), Freedom of Speech, 1st ed., Routledge, London, 2018. doi:10.4324/9781315181981-3
- [21] S. Frier, No Filter: The Inside Story of Instagram, Simon & Schuster, New York, NY, USA, 2020.
- [22] C. Lewis, Irresistible Apps: Motivational Design Patterns for Apps, Games, and Web-based Communities, Apress, New York, NY, USA, 2014.
- [23] Competition and Markets Authority, Online Choice Architecture: How digital design can harm competition and consumers, 2022. URL: https://assets.publishing.service.gov.uk/media/624c27c68fa8f527710aaf58/Online_choice_architecture_discussion_paper.pdf
- [24] OECD, Dark commercial patterns, 2022. URL: <https://www.oecd.org/digital/dark-commercial-patterns-44f5e846-en.htm>
- [25] M. Potel-Saville, M. Da Rocha, From Dark Patterns to Fair Patterns? Usable Taxonomy to Contribute Solving the Issue with Countermeasures, Annual Privacy Forum, Lyon, France, 2023.
- [26] D. Totzek, G. Jurgensen, Many a little makes a mickle: Why do consumers negatively react to sequential price disclosure?, Psychology and Marketing 38 (2020) 113-129. doi:10.1002/mar.21426
- [27] C. Voigt, S. Schlögl, A. Groth, Dark Patterns in Online Shopping: of Sneaky Tricks, Perceived Annoyance and Respective Brand Trust, in: Proceedings of the 23rd HCI International Conference, HCII 2021, Springer, Cham, 2021, pp. 143-155. doi:10.1007/978-3-030-77750-0_10
- [28] C. M. Gray, C. Santos, N. Bielova, Towards a Preliminary Ontology of Dark Patterns Knowledge, in: Extended Abstracts of the 2023 CHI Conference on Human Factors in

- Computing Systems, CHI EA '23, ACM, New York, NY, USA, 2023. doi:10.1145/3544549.3585676
- [29] D. Berdichevsky, E. Neuwander, Toward an ethics of persuasive technology, *Communications of the ACM* 42 (1999) 51–58. doi:10.1145/301353.301410
- [30] Michael Nagenborg, Surveillance and persuasion, *Ethics and Information Technology* 16 (2024) 43–49. doi:10.1007/s10676-014-9339-4
- [31] D. Susser, B. Roessler, H. Nissenbaum, Technology, autonomy, and manipulation, *Internet Policy Review* 8 (2019). doi:10.14763/2019.2.1410
- [32] H. Latham, 'Ethical design' is a dangerous term: How to approach being ethical, *Medium*, 2018. URL: <https://uxdesign.cc/ethical-design-is-a-dangerous-term-b314a5e385f4>
- [33] M. Maier, R. Harr, Dark Design Patterns: An End-User Perspective, *Human Technology* 16 (2020) 170-199. doi:10.17011/ht/urn.202008245641
- [34] S. Buss, A. Westlund, Personal Autonomy, in: E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, Spring 2018 Edition, 2018. URL: <https://plato.stanford.edu/archives/spr2018/entries/personal-autonomy/>
- [35] A. Roskies, Neuroethics, in: E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, Spring 2021 Edition, 2021. URL: <https://plato.stanford.edu/archives/spr2021/entries/neuroethics/>
- [36] A. Vugts, M. van den Hoven, E. de Vet, M. Verweij, How autonomy is understood in discussions on the ethics of nudging, *Behavioural Public Policy* 4 (2020) 108-123. doi:10.1017/bpp.2018.5
- [37] S. Ahuja, J. Kumar, Assistant or Master: Envisioning the User Autonomy Implications of Virtual Assistants, in: *Proceedings of the 4th Conference on Conversational User Interfaces, CUI '22*, Association for Computing Machinery, New York, NY, USA, 2022, Article 4, pp. 1–5. doi:10.1145/3543829.3544514
- [38] O. Friedrich, E. Racine, S. Steinert, J. Pömsl, R. J. Jox, An analysis of the impact of brain-computer interfaces on autonomy, *Neuroethics* 14 (2018) 17-29. doi:10.1007/s12152-018-9364-9
- [39] P. G. Hansen, A. M. Jespersen, Nudge and the Manipulation of Choice: A Framework for the Responsible Use of the Nudge Approach to Behaviour Change in Public Policy, *European Journal of Risk Regulation* 1 (2013) 3-28. URL: <https://ssrn.com/abstract=2555337>
- [40] J. W. Brehm, *A theory of psychological reactance*, Academic Press, 1966.
- [41] R. H. Thaler, C. R. Sunstein, *Nudge*, Yale University Press, New Haven, USA, 2008.
- [42] M. Betzler, Authenticity and self-governance, in: M. Salmela, V. Mayer (eds.), *Emotions, ethics, and authenticity*, John Benjamins Publishing Company, Amsterdam, The Netherlands, 2009.
- [43] N. Levy, *Neuroethics*, Cambridge University Press, Cambridge, United Kingdom, 2007.
- [44] D. A. Norman, Cognitive engineering, in: D. A. Norman, S. W. Draper (eds.), *User centered system design, New Perspectives on Human-Computer Interaction*, Lawrence Erlbaum Associates, Hillsdale, NJ, USA, 1986, pp. 31-65.

- [45] G. D. Abowd, R. Beale, Users, systems and interfaces: A unifying framework for interaction, in: Proceedings of the HCI '91 Conference, Cambridge University Press, Cambridge, United Kingdom, 1991, pp. 73-87.
- [46] B. Moggridge, Designing Interactions, MIT press, Cambridge, Massachusetts, USA, 2006.
- [47] K. Silver, What puts the design in interaction design, UX Matters, 2007. URL: <https://www.uxmatters.com/mt/archives/2007/07/what-puts-the-design-in-interaction-design.php>
- [48] P. Morville, User experience design, 2004. URL: https://semanticstudios.com/user_experience_design/
- [49] J. J. Garrett, The Elements of User Experience: User-Centered Design for the Web and Beyond, Pearson Education, Berkeley, CA, USA, 2010.
- [50] S. Michie, M. M. van Stralen, R. West, The behaviour change wheel: a new method for characterising and designing behaviour change interventions, Implementation Science 6 (2011) 1-12. doi:10.1186/1748-5908-6-42
- [51] N. Tromp, P. Hekkert, P. P. Verbeek, Design for Socially Responsible Behavior: A Classification of Influence Based on Intended User Experience, Design Issues 27 (2011) 3-19. URL: <http://www.jstor.org/stable/41261940>
- [52] C. Meske, T. Potthoff, The DINU-model—A Process Model for the Design of Nudges, in: Proceedings of the 25th European Conference on Information Systems (ECIS), Guimarães, Portugal, 2017.
- [53] M. Jesse, D. Jannach, Digital nudging with recommender systems: Survey and future directions, Computers in Human Behavior Reports 3 (2021) 100052. doi:10.1016/j.chbr.2020.100052
- [54] R. Münscher, M. Vetter, T. Scheuerle, A review and taxonomy of choice architecture techniques, Journal of Behavioral Decision Making 29 (2016) 511-524. doi:10.1002/bdm.1897
- [55] D. C. Mohr, S. M. Schueller, E. Montague, M. N. Burns, P. Rashidi, The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions, Journal of Medical Internet Research 16 (2014). doi:10.2196/jmir.3077
- [56] L. Yardley, B. J. Spring, H. Riper, L. G. Morrison, D. H. Crane, K. Curtis, G. C. Merchant, F. Naughton, A. Blandford, Understanding and promoting effective engagement with digital behavior change interventions, American Journal of Preventive Medicine 51 (2016) 833-842. doi:10.1016/j.amepre.2016.06.015
- [57] T. Sharon, Self-Tracking for Health and the Quantified Self: Re-Articulating Autonomy, Solidarity, and Authenticity in an Age of Personalized Healthcare, Philosophy & Technology 30 (2017) 93-121. doi:10.1007/s13347-016-0215-5
- [58] E. Pariser, The filter bubble: What the Internet is hiding from you, Penguin Books, London, United Kingdom, 2011.

A. Appendix

Table A.1

Coding Guide for 5 Layers of UX Design (based on [49])

Code	Coding Guide
Surface	This layer consists of visual design elements, i.e., the look of a finished product.
Skeleton	This layer consists of interface design, i.e., the arrangement of visual design elements. It also consists of information design, i.e., the presentation of information in a way that facilitates understanding.
Structure	This layer consists of interaction design, i.e., how the system behaves in response to a user. It also consists of information architecture, i.e., the arrangement of informational elements within the product space.
Scope	This layer consists of functional specifications, i.e., a detailed description of the feature set of a product. It also consists of the content requirements of a product.
Strategy	This layer consists of user needs, i.e., what the audience wants from a product. It also consists of product objectives defined by business goals.

Table A.2

Coding Guide for Behavior Design Model (based on [2])

Code	Coding Guide
Motivation	This factor affects users' desire or intent to perform a behavior. Motivational types can include pleasure / pain, hope / fear, social acceptance / rejection, etc.
Ability	This factor affects users' ability or capability to perform a behavior. Barriers to ability can include time, money, physical effort, mental effort, and breaking routine.
Triggers	This factor provides users with an opportunity to perform a behavior through prompts, calls to action, cues, etc. It tells people to perform a behavior now.