UX Design: the Impacts on Physiological Responses

Nataša Miletić[1], dr. Matjaž Kljun[2] and dr. Klen Čopič Pucihar[2]

[1]University of Primorska, The Faculty of Mathematics, Natural Sciences and Information Technologies, Glagoljaška ulica 8, 6000 Koper

Abstract

Dark patterns are deceptive design techniques used to trick and manipulate users into taking actions they did not initially intend, often exploiting cognitive biases and obscuring true choices to benefit the designer or service provider. The study described is an initial investigation into impacts of dark patterns on physiological responses during interactive tasks. We used a combination of questionnaires and biometric data to asses user frustration, engagement, and emotional responses to various dark patterns. We measured anger, disgust, joy, and engagement through facial expression analysis as well ad galvanic skin response (GSR). The research seeks to contribute to further understanding of the implications of dark patterns on users' physiological responses.

Keywords

dark patterns, physiological responses, UX design

1. Introduction

1.1. Problem Statement

Conventional web design and development practices often incorporate user experience (UX) techniques aimed at enhancing usability and satisfaction. However, some techniques, known as dark patterns, intentionally manipulate users into taking actions they might not otherwise take [**?**]. Dark patterns can include misleading information, forced continuity, and hidden costs, among others. While experienced users may recognize and avoid these tactics, novice users are more susceptible to them, which can lead to frustration, mistrust, and ethical concerns [**?**]. Understanding how dark patterns affect users, such as stress levels and eye movement, can provide insights into the immediate impact of these practices on user engagement and well being. Recent research highlighted the need for more user-centered design approaches that prioritize user welfare over manipulative tactics. As such, there is a need to better understand the impact of dark patterns, combining both subjective experiences and objective physiological data.

1.2. Research Questions and Objectives

In this study, we tried to answer the following research questions (RQ):

HCI SI 2024: Human-Computer Interaction Slovenia 2024, November 8th, 2024, Ljubljana, Slovenia

89191146@student.upr.si (N. Miletić); matjaz.kljun@upr.si (dr. M. Kljun); klen.copic@famnit.upr.si (dr. K. Č. Pucihar)

RQ1 Will the presence of dark patterns in UX design have an impact on physiological responses?
RQ2 Will the presence of different types of dark patterns have a different influence on different users?

To this end we have run two user studies. In Study A participants completed a questionnaire about their experiences with and perceptions of dark patterns, specifically which ones they find most annoying. In Study B participants navigated a different web pages designed to feature a variety of dark patterns. During Study B we captured eye-tracking, GSR and heart rate, and participants had to fill in a couple of questionnaires.

2. Study A

2.1. Method

To investigate user perceptions and annoyance levels associated with various dark patterns, a structured questionnaire was developed and made available online. The questionnaire included 11 distinct types of dark patterns selected based on its common occurrence in digital interfaces and in the academic literature [**?**]. Dark patterns included in the study are:

- **Disguised Ads:** Making advertisements look like part of the page content.
- **Social Pyramid:** Convincing users to share contacts that will be used for marketing purposes.
- **Scaremongering or Toying with emotion:** Making users believe they are in danger.
- **Deceptive Wording:** Using complex wording to obscure the final action as a subset of the pattern called hidden information or aesthetic manipulation.
- **Sneak into Basket:** Convincing users to add more products to their basket to increase spending.
- **Forced Action or Pre-selection:** Offering pre-selected expensive or firm friendly options first.
- **Price Comparison Prevention:** Combining prices in a complex manner to prevent easy comparison of products.
- **Bait and Switch:** Offering an apparent bargain with the intention of substituting inferior or more expensive goods.
- **Confirmshaming or "FOMO" (Fear of Missing Out):** Influencing users' decisions by triggering uncomfortable emotions about missing out.
- **Hidden Costs:** Not showing the full price upfront.
- **Privacy Zuckering:** Tricking users into sharing more information about themselves than they would otherwise.

The study involved 39 participants that were regular users of digital interfaces, encompassing a wide range of ages, professions, and levels of technical proficiency recruited through social media invitations. Participants were presented with a detailed description and a visual representation of each dark pattern as seen in **??**. The visual aids were designed to mimic real-life scenarios in

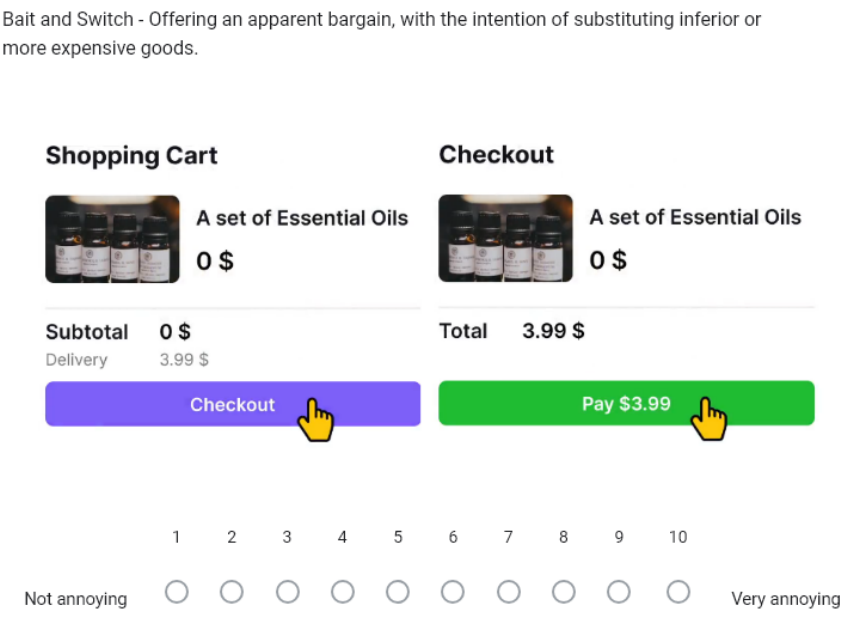

Figure 1: 'Bait and Switch' visual representation

which these dark patterns commonly appear, thereby providing participants with a realistic context for their ratings.

Each participant was asked to rate their level of annoyance for each dark pattern on a scale from 1 to 10, where 1 indicated "not annoying" and 10 indicated "very annoying". All questions were present on one page so participants could change their answers anytime before submitting the form. They were informed beforehand about the nature of the questions and the purpose of their answers. Additionally, they were provided with contact information in case they were interested in the final results of the study.

2.2. Results

Participants in the study faced no time constraints and were able to complete the questionnaire at their own pace. The responses were based on their prior interactions with various dark patterns encountered during their web usage. The results reflect a range of experiences. The descriptive statistics for the ratings of the 11 dark patterns are presented in **??**. *Privacy Zuckering* received the highest average annoyance rating ($\bar{x} = 8.92$, $SD = 1.88$), indicating a strong consensus among participants regarding its irritation. Similarly, *Hidden Costs* was rated highly annoying ($\bar{x} = 8.85$, $SD = 1.93$). In contrast *Sneak into Basket* ($\bar{x} = 6.62$, $SD = 2.60$) and *Confirmshaming or FOMO* ($\bar{x} = 6.72$, $SD = 2.74$) were rated lower on average, suggesting they are perceived as less annoying or not encountered often. However, the higher standard deviations for these patterns indicate a wider range of opinions, making them less universally disliked.

Table 1
Descriptive statistics of dark patterns from Study A

Dark Pattern	Mean	Median	Standard Deviation	Minimum	Maximum
Disguised Ads	6.92	7.00	2.64	1	10
Social Proof Spoofing	8.18	9.00	2.37	1	10
Scaremongering	8.28	10.00	2.74	1	10
Deceptive Wording	8.31	9.00	1.95	3	10
Sneak into Basket	6.62	7.00	2.60	1	10
Forced Action	6.77	7.00	2.69	1	10
Price Comparison Prevention	6.97	7.00	2.73	1	10
Bait and Switch	7.15	7.00	2.68	1	10
Confirmshaming or FOMO	6.72	7.00	2.74	1	10
Hidden Costs	8.85	10.00	1.93	2	10
Privacy Zuckering	8.92	10.00	1.88	3	10

3. Study B

3.1. Method

In this study we conducted a controlled experiment to measure physiological responses of participants exposed to various dark patterns. Users had to complete 5 tasks on different web pages. After completing a predefined task on each web page, participant answered the NASA Task Load Index (NASA-TLX)[1], which measures the perceived workload in order to assess a task performance. The instructions given for each task were intentionally minimal to replicate a realistic user experience. No additional information or hints were provided, ensuring that participants relied solely on their skills to navigate, search, compare products through potentially misleading or irrelevant data, provided they remained on the same website. This allowed us to evaluate participant responses to realistic examples of dark patterns.

This study involved nine participants, all students from different fields of study (computer science, biopsychology, management) and aged between 21 and 24 years ($\bar{x} = 22.88$, $SD = 1.09$). These participants were regular computer users, ensuring that the sample was representative of typical student user. Participants were fully informed about studies' objectives, provided with consent forms, and assured of their right to withdraw at any time without repercussions. Participants were briefed on the general purpose of the study, though specific details about the dark patterns or expected outcomes were not disclosed to prevent any bias in their interactions. Contact information was provided for participants interested in the final results.

3.2. Tasks

The study consisted of five distinct tasks, each needed to be completed on an existing web page. We used existing web pages to simulate real-life scenarios, ensuring that the participants' reactions would be genuine and reflective of their typical user behavior. Some of the web pages were also hosting illegitimate content but were selected due to the high number of dark patterns used.

[1]NASA-TLX https://en.wikipedia.org/wiki/NASA-TLX

- **Task A:** Participants were required to navigate APK Mirror[2], website hosting downloads of various software Android application packages (APKs). Participants were instructed to locate and download a specific version of the YouTube application (19.16.39). This task assessed their ability to discern legitimate links from disguised ads and irrelevant results.
- **Task B:** Participants were required to navigate to DaddyLive[3], a website streaming various sports events and TV channels. The specific objective was to find and stream a particular channel, Arena Sport Serbia, and ensure the video was fullscreened and unmuted. This task focused on dealing with a stream of pop-up windows.
- **Task C:** The task required participants to navigate the Amazon UK[4],website to find and identify the cheapest physical copy of the book "Harry Potter and the Order of the Phoenix" (either paperback or hardcover) in any language. This task tested their proficiency in using e-commerce filters and search functions among misleading listings.
- **Task D:** The task involved navigating The Pirate Bay[5] torrent site, finding and downloading a torrent file for participants' favorite video game. This task aimed to assess the ability to identify the correct download link among various misleading elements and gather data on user reactions to site responsiveness and redirect behaviors.
- **Task E:** The fifth task involved navigating the MovieWatcher[6] website, where participants were asked to locate and play any of the movies from "The Lord of the Rings" trilogy. This task was designed to evaluate their persistence and time to abandonment when faced with often unresponsive and at times non-functional website.

3.3. Measured Parameters

Several physiological parameters were measured to assess the participants' responses besides the NASA-TLX perceived workload questionnaire. We used the iMotions software to capture the following data:

- Eye Tracking: Tobii Pro Spectrum screen was used to capture eye tracking data including pupil dilation.
- Heartbeat and Galvanic Skin Response: Shimer3 GSR+ sensor on participant's wrist captured the heartbeat and skin conductivity as indicators of physiological stress.
- Facial Emotion Analysis: High-resolution camera recorded participants' facial expressions with Affectiva AFFDEX [**?**] embedded in iMotions software to detect emotional positive and negative responses.

It needs to be noted that we did not compare the group exposed to dark patterns with a group not exposed to them and so we do not use comparative statistics. For facial expression we aggregated the data for negative and positive emotions as in [**?**]. From the eye tracking data we focused in particular on **Pupil Dilation** [**?**]. Pupil dilation refers to the enlargement

[2]APK Mirror https://www.apkmirror.com/
[3]DaddyLive https://daddylive.watch/
[4]Amazon UK https://www.amazon.co.uk/
[5]The Pirate Bay https://thepiratebay.org/
[6]MovieWatcher https://moviewatcher-to.lol/

of the pupils and can provide insights into attention, interest, emotion, mental workload, and arousal. For NASA-TLX we used the **Wilcoxon signed-rank** [**?**] test to compare subjective workload scores between different tasks.

3.4. Results

In Task A (APK Mirror) participants often clicked on misleading ads disguised as download buttons, leading to frustration and increased time spent on the task. The cognitive load was higher (see **??**), possibly due to constant re-evaluation of links, making this task one of the more mentally demanding. In Task B (DaddyLive) the pop-up-heavy nature of the website introduced forced action patterns, where users had to close numerous pop-ups before accessing the desired stream. This task caused significant annoyance and distraction, especially when participants struggled to return to the main content. Users showed heightened levels of heartbeat (see **??**) due to the persistence of pop-ups, but once they managed to complete the task, the frustration diminished quickly.

Task C (Amazon UK) focused on price comparison and search filters. Participants faced obfuscation through misleading listings and non-relevant search results. However, the familiar interface of Amazon helped mitigate extreme frustration, and participants managed to complete the task with moderate cognitive load. The emotional responses here were intense with positive emotions prevailing. Some participants reported being overwhelmed by the sheer number of listings.

Task D (Pirate Bay Torrent Search) made participants face false links and numerous redirects, that caused significant frustration and most negative responses across tasks. Users struggled to find the correct download link amidst misleading ads, leading to heightened stress levels. This task was particularly demanding due to the site's aggressive dark patterns, including constant redirects and fake download buttons, which resulted in a longer task completion time.

Task E (MovieWatcher Website): This task proved to be the most frustrating due to completely unresponsive and non-functional elements. The MovieWatcher website was riddled with dead ends and forced continuity, as users were bombarded with ads and reloading pages that prevented them from completing the task. Most participants abandoned the task, showcasing its high emotional toll.

Overall, Task E was reportedly very frustrating due to severe usability issues and dark patterns that rendered the website practically unusable. Task D and B followed closely, with deceptive elements frustrating users to a high degree. Task B, though disruptive, was manageable once users navigated past the pop-ups, while Task C posed moderate challenges with less emotional strain. Task A, though complex, did not elicit the same level of frustration as tasks that involved more aggressive dark patterns like those in Task D and Task E.

3.4.1. NASA-TLX

Task E (MovieWatcher) was consistently perceived as more demanding than Tasks A (APKMirror), and Task C (Amazon), which could suggest that Task E involves higher cognitive or physical demands. This may be attributed to Task E's complexity or the specific skills required, leading to increased mental or physical strain compared to the other tasks. Task B and Task E have

Table 2
Mean (Standard Deviation) for each task from Study B

Task	Negative Emotions in % (SD)	Positive Emotions in % (SD)	Heart Rate	Pupil Dilation
A	27.21 (10.04)	10.90 (4.88)	97.81 (7.55)	2.96 (0.41)
B	24.67 (8.07)	17.62 (4.90)	72.39 (12.91)	3.52 (1.12)
C	41.65 (12.51)	69.84 (11.98)	88.57 (10.65)	2.89 (1.46)
D	59.42 (9.42)	12.64 (8.10)	89.65 (12.49)	2.24 (0.43)
E	28.10 (10.60)	14.05 (6.27)	101.99 (34.89)	4.19 (1.96)

similar perceived workloads. This could be relevant for acknowledging that tasks are designed to be of similar difficulty or require similar skill sets.

Frustration ratings show an upward trend across tasks, reaching higher levels in later tasks. As tasks become more difficult, participants experience greater frustration. This suggests that the increasing complexity or challenges of the tasks contribute to a heightened sense of frustration.

Temporal demand ratings show a rising trend from Task A through Task E, indicating an increasing sense of urgency or pressure to complete tasks. This reflects that participants feel more pressured to manage their time effectively as they progress through the tasks.

The overall performance ratings across tasks tend to cluster around the 50 percentage mark. This consistent midpoint suggests that participants generally felt they performed at an average level, regardless of the task's complexity or difficulty.

Table 3
Mean ratings for each measure across tasks from study B

Measure	Task A	Task B	Task C	Task D	Task E
Mental Demand	26.67	47.78	29.44	37.78	43.33
Physical Demand	11.11	18.89	20.56	27.78	32.78
Temporal Demand	30.56	46.67	36.11	51.11	63.33
Performance	48.89	45.00	57.22	51.11	47.78
Effort	25.00	55.00	35.00	52.22	62.22
Frustration	32.22	68.89	32.78	60.00	81.11
Overall	28.24	53.89	35.28	51.28	54.89

4. Discussion

RQ1 Will the presence of dark patterns in UX design have an impact on physiological responses, such as stress levels and pulse?

During the experiments, we observed significant fluctuations in GSR conductance levels, particularly when users encountered dark patterns. These patterns often involve misleading or obstructive elements in the user interface designed to manipulate user behavior, such as repetitive pop-up ads or misleading navigation elements. Participants exhibited notable increases in GSR conductance when they struggled to complete tasks on the first attempt. For instance,

users who had to repeatedly close pop-up ads experienced elevated stress levels. The need to frequently interact with these deceptive elements caused frustration, which was reflected in their physiological responses. Waiting for pages to load or when faced with delays in the interface was also accompanied by increases in GSR conductance, suggesting that waiting and delays contributed to heightened stress and decreased user satisfaction.

Participants were observed to smile even when they were visibly frustrated, such as when they could not locate a desired button or feature. This reaction underscores a complex interplay between emotional responses and physiological arousal. The incongruity between facial expressions and physiological data highlights how users' external behavior may not always align with their internal stress levels.

We noticed that participants often fixed their gaze on the upper right corner of the window, anticipating that a pop-up ad would have an 'X' button for closing. This fixation point was a behavioral indicator of their expectation and frustration with the dark patterns in the interface.

RQ2 Will the presence of different types of dark patterns have a different influence on different users?

The findings indicate that the impact of dark patterns on users varies significantly depending on the nature of the task and users' experience with the web and existing dark patterns. Tasks involving complex interactions or repetitive attempts revealed more pronounced physiological and emotional responses. A notable observation was that many participants smiled while expressing frustration when repeatedly encountering various pop-up ads. This reaction could be attributed to the lab setting, where participants might have been more conscious of their behavior or felt a heightened sense of self-awareness compared to their typical at-home environment.

These variations highlight the need for a nuanced understanding of how different dark patterns affect user behavior, underscoring the importance of continued research into their effects on user trust and stress.

5. Limitations and Future Work

We are not comparing the use of dark patterns to any baseline where there are no dark patterns. So we report only on the responses for particular tasks completed on particular web sites. Further investigation should provide a baseline and compare it to the design with dark patterns. Due to time constraints, our study was conducted with a relatively small sample size. All participants were constant computer users, which may have introduced a bias in their responses, as their habitual interaction with technology could influence their reactions to dark patterns differently compared to occasional users. Additionally, the controlled lab environment may not fully capture the complexities of users' interactions with dark patterns in their natural settings.

6. Conclusion

Given the ubiquity of dark patterns on the web [**?**] we run two user studies exploring attitudes towards dark patterns and their physiological responses when encountering them. Study

A included a questionnaire of 11 dark patterns and asked participants how annoying was a particular pattern on a scale from 1 (not at all) to 10 (very annoying). In study B we asked participants to complete 5 tasks, each on a different website featuring a variety of dark patterns

Through selected tasks, we observe distinct variations in users responses to different types of dark patterns, underscoring the complex interplay between design elements and user behavior. The results highlight that misleading design practices can exaggerate user frustration and stress.

The task/specific results highlighted how users frustration and stress were particularly pronounced in scenarios involving repeated obstacles or misleading interface elements, such as those encountered during searches or interactions with unresponsive websites. These finding emphasize that dark patterns significantly disrupt user experience, and increase cognitive load.

References

[] Dark patterns: Deceptive user interfaces, https://www.darkpatterns.org, 2010. Accessed: 2024-09-30.

[] 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, 2018, pp. 1–14.

[] J. Luguri, L. J. Strahilevitz, Shining a light on dark patterns, Journal of Legal Analysis 13 (2021) 43–109.

[] M. Bishay, K. Preston, M. Strafuss, G. Page, J. Turcot, M. Mavadati, Affdex 2.0: A real-time facial expression analysis toolkit, in: 2023 IEEE 17th international conference on automatic face and gesture recognition (FG), IEEE, 2023, pp. 1–8.

[] M. Weerasinghe, K. Čopič Pucihar, J. Ducasse, A. Quigley, A. Toniolo, A. Miguel, N. Caluya, M. Kljun, Exploring the future building: representational effects on projecting oneself into the future office space, Virtual Reality 27 (2023) 51–70.

[] Wikipedia, Mydriasis wikipedia, https://en.wikipedia.org/wiki/Mydriasis, 2024.

[] Datab.net, Wilcoxon signed rank test, https://datatab.net/tutorial/wilcoxon-test, 2024.

[] A. Mathur, K. M., Dark patterns at scale: Findings from a crawl of 11,000 shopping websites, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, ACM, 2019, pp. 1–12.