



Faculty 3: Mathematics and Computer Science  
Study Program: Digital Media (B.Sc.)

BACHELOR THESIS

# Complicity

A Game-Based Study of How Role Perspective Shapes Awareness and  
Understanding of Exploitative Design Practices

**Katharina Rex**

Matrikelnummer: 6188763

Datum: May 25, 2026

Supervisor 1: Prof. Dr. Rainer Malaka  
Supervisor 2: Dr. René Weller  
Advisor: Dr.-Ing. Nima Zargham

## Abstract

Dark Patterns shape user behavior through interface designs that obscure choices, nudge decisions, or restrict agency. Existing research predominantly examines these practices from the user standpoint, focusing on their detection, classification and ethical implications.

This study presents *Complicity*, a first-person 3D game designed to enable a comparative analysis of role perspectives in a game environment centered on Dark Patterns. Using a between-subjects experimental design, participants were randomly assigned to one of two conditions: a Victim role, navigating a series of playful mini-games that contain Dark Patterns, or an Exploiter role, deploying these patterns to achieve simulated performance goals. The study investigates whether embodying the Exploiter affects immediate recognition of manipulative design strategies and supports longer-term awareness of Dark Patterns more strongly than the Victim condition. Quantitative data was collected through pre- and post-intervention awareness tests, while qualitative insights from post-study semi-structured interviews provide a deeper understanding of participants' reasoning and their reflections on the experience.

The findings indicate that the intervention increased participants' ability to recognize manipulative interface strategies across both conditions. While both roles successfully fostered awareness, the Exploiter condition demonstrated stronger tendencies toward critical reflection and analytical understanding of how and why Dark Patterns are implemented. In contrast, the Victim condition primarily increased awareness through direct exposure to Dark Patterns and emotional frustration. Overall, the study suggests that the game-based experience and the perspective shift toward the designer role can serve as an effective educational approach for raising awareness of manipulative design practices.

# Contents

<b>1</b>	<b>INTRODUCTION &amp; MOTIVATION</b>	<b>4</b>
<b>2</b>	<b>RELATED WORK</b>	<b>5</b>
2.1	Identifying Dark Patterns and User Blindness . . . . .	5
2.2	Existing Awareness Tools . . . . .	5
2.3	Game-Based Learning and Role Perspectives . . . . .	6
2.4	Moral Disengagement and Reflective Roleplay . . . . .	7
<b>3</b>	<b>GAME DESIGN</b>	<b>7</b>
3.1	Victim Role . . . . .	8
3.2	Exploiter Role . . . . .	8
3.3	Task Structure and Dark Pattern Scenarios . . . . .	9
3.4	Game World and Aesthetic Framing . . . . .	14
<b>4</b>	<b>METHODOLOGY</b>	<b>14</b>
4.1	Dark Pattern Detection Test . . . . .	14
4.2	Post-Intervention Questionnaires . . . . .	17
4.3	Follow-Up Survey . . . . .	17
4.4	Study Procedure . . . . .	18
4.5	Pilot Testing . . . . .	19
4.6	Participants . . . . .	19
4.7	Pre-Intervention Group Comparison . . . . .	19
<b>5</b>	<b>FINDINGS</b>	<b>20</b>
5.1	Dark Pattern Detection Test - Results . . . . .	20
5.2	Self-Reported Awareness and Reflection . . . . .	25
5.3	Player Experience Inventory (PXI) . . . . .	26
5.4	Behavioral Data and Log File Analysis . . . . .	28
5.5	Qualitative Findings . . . . .	29
5.6	Follow-up Survey - Results . . . . .	32
<b>6</b>	<b>DISCUSSION</b>	<b>35</b>
6.1	Evaluation of the Dark Pattern Detection Tests . . . . .	35

6.2	RQ 1: The Intervention as an Educational Tool . . . . .	37
6.3	RQ 2: Influence of Role Perspectives on Immediate Awareness . . . . .	38
6.4	RQ 3: Sustained Vigilance and Long-term Impact . . . . .	39
6.5	Practical Implications . . . . .	40
<b>7</b>	<b>LIMITATIONS &amp; FUTURE WORK</b>	<b>41</b>
<b>8</b>	<b>CONCLUSION</b>	<b>42</b>

# 1 INTRODUCTION & MOTIVATION

While interface design aims to support usability and user goals, economic objectives such as increasing engagement and revenue are often prioritized. Therefore, interface elements may not only support but actively steer user behavior. [1] Within this environment, Dark Patterns have emerged as a prominent form of manipulative design. Dark Patterns refer to interface design choices that push users toward actions that primarily benefit the service provider, often by obscuring alternatives, creating misleading choices, or exploiting cognitive biases. [2, 3] These practices are widespread across websites, mobile applications and games and have become a recurring characteristic of everyday digital interaction. [4]

This study is motivated by an initial engagement with Shoshana Zuboffs’ concept of “Surveillance Capitalism” [5], in which human experience is treated as raw material for behavioral data. This data is further processed into predictive models designed not only to anticipate but also to influence user behavior and is traded to guarantee outcomes for advertisers. To support these systems, digital environments incorporate mechanisms that guide users toward economically desirable actions. Dark Patterns serve as a primary tool for this ‘behavioral modification’, ensuring that users remain within digital ecosystems and continue to provide the data necessary for profit. Although public awareness and legal regulations have increased, these patterns persist, as they are embedded in the economic logic of modern digital platforms. The augmented reality game Pokémon Go is a well-known example of this dynamic, as it encouraged players to visit sponsored locations through gameplay without making the commercial motives immediately obvious. [6, 5]

Despite their prevalence, users often struggle to detect the manipulative intent behind these interfaces. Research describes this as “dark pattern blindness” (or DP-blindness), a phenomenon where users may feel frustrated but fail to recognize the specific manipulative strategies used against them. [7] In response, existing research and educational efforts have mainly focused on the user’s perspective, particularly on recognizing Dark Patterns, classifying them, protecting consumers, and promoting ethical design practices. While this work establishes awareness and shared terminology, it primarily frames users as passive recipients of manipulation. However, it remains unclear, whether directly engaging with these practices from a design perspective might influence awareness differently and if taking on the role of the “exploiter” and actively implementing Dark Patterns within a constrained system, can potentially promote a deeper understanding than experiencing them as a user.

This study uses a custom-developed first-person 3D game Complicity, to investigate how different role perspectives influence the understanding of manipulative design. The game features two conditions: in one, players take on the role of a user navigating and experiencing Dark Patterns; in the other, they assume the role of a UI/UX designer tasked with implementing them to meet corporate goals. This research examines whether playful participation in the design process, combined with the confrontation of its consequences, fosters a different or more sustained form of awareness compared to user-centered approaches.

The study is guided by the following three research questions:

- **RQ 1:** How does embodying the Exploiter- or Victim role in a first-person interactive narrative game influence player’s immediate awareness of manipulative design practices?
- **RQ 2:** Does embodying the Exploiter role (actively deploying manipulative design and experiencing the consequences) lead to stronger or deeper awareness of Dark Patterns compared to the Victim role?
- **RQ 3:** Does this game-based experience foster long-term awareness of manipulative design practices and does one role sustain that awareness more effectively than the other?

By addressing these questions, the study explores how role-based interaction shapes awareness of deceptive design practices, with particular attention to how participants reflect on their own design decisions within system constraints.

## 2 RELATED WORK

This section situates the study within existing research on Dark Patterns, user awareness, and role-based learning. It examines game-based learning and role-taking as approaches to studying awareness, and introduces relevant psychological frameworks that inform the design of this study.

### 2.1 Identifying Dark Patterns and User Blindness

To create a consistent vocabulary for research, scholars have developed detailed lists and ontologies of known Dark Patterns. For this study, the Ontology of Dark Patterns Knowledge proposed by Gray et al. [4] was used as the primary reference, because it provides a structured way to understand how different patterns relate to one another and how they manifest across different platforms. From this ontology, specific patterns were selected based on their translatability into interactive game mechanics for both the Victim and Exploiter roles. The chosen patterns Confirmshaming, Roach Motel and Sneak into Basket, were selected because they are commonly encountered and often leave a lasting impression on users due to their frustrating nature. Confirmshaming uses emotionally manipulative language to steer choices, while the Roach Motel pattern focuses on creating structural friction, making it easy for users to sign up for a service but deliberately difficult to cancel. The Sneak into Basket pattern was included due to its significant potential for financial harm. Although such practices are restricted under EU consumer protection laws, variants of this pattern continue to appear in practice, often in more subtle forms such as pre-selected extras or automatically included services like insurances. [7] This suggests that manipulative strategies may persist in forms that are less explicit, even where more overt implementations are legally restricted.

While Dark Patterns are the primary means of steering behavior, they are often difficult for users to recognize. This is frequently described as "dark pattern blindness," [7] where users may feel frustrated during an interaction but fail to identify the specific manipulative strategy being used. [8] The phenomenon of user blindness makes manipulative designs particularly dangerous. In a large-scale online experiment, Luguri and Strahilevitz demonstrated that Dark Patterns, even mild ones, are highly effective at coercing user behavior. They found that users exposed to these patterns were significantly more likely to accept offers they would have otherwise rejected. [9] This highlights that merely encountering a Dark Pattern does not equip a user to recognize or resist it, making 'dark pattern blindness' a critical vulnerability that designers can exploit. Another investigation into popular mobile applications found that while 95% of the sampled apps contained at least one Dark Pattern, users were often unable to perceive them. Specifically, results showed that 55% of participants failed to spot malicious elements and another 20% remained unsure of their presence. Only 25% of users were able to correctly identify a Dark Pattern when present. However, the same study noted that 86% of users could accurately recognize the absence of Dark Patterns in a control task. [7] This suggests that while users are generally capable of recognizing "fair" interfaces, they lack the necessary awareness to detect active deception, highlighting the need for more effective educational interventions. Furthermore, even when users are generally aware that platforms use manipulative designs, this awareness does not necessarily equip them to oppose their influence. Because these interfaces exploit cognitive biases, users often recognize that a design feels manipulative but remain unsure of the specific mechanisms being used against them or the actual harm they might suffer. [10] This research examines if shifting the role perspective can provide a clearer insight into how these design strategies are used.

### 2.2 Existing Awareness Tools

Several research projects have developed games to raise awareness about deceptive design. Kronhardt et al. (2024) introduced "Trickery", a narrative serious game that incorporates seven gamified deceptive patterns. [11] The game raises awareness by making players experience the direct consequences of being manipulated as they attempt to navigate and escape a laboratory environment controlled by an antagonistic scientist. Other researchers have explored different approaches, such as Meinhardt et al. (2025) in their study "Mind Games!" [6], which examined the impact of Dark

Patterns in Mixed Reality (MR) scenarios. By simulating a city walk through videos, they demonstrated that Dark Patterns in MR, such as forced registration or information hiding, significantly reduce user comfort.

While these approaches provide insight into user experience and emotional response, they focus primarily on the user as the recipient of manipulation. In addition to these Dark Pattern-specific tools, the present study is influenced by the narrative mechanics of the video game “Papers, Please” [12], which demonstrates how systemic and psychological pressure can be used to force a player into making morally questionable decisions. This mechanic is reflected in the "Exploiter" role of the current study, where players must navigate professional constraints and corporate goals that often conflict with ethical design. Such environments can place UX practitioners in situations where business interests are prioritized over user well-being. This tension relates to the concept of “role morality,” where individuals adopt different moral standards within professional contexts and may distance themselves from the ethical consequences of their decisions. [13] The importance of a game’s theme is supported by Grudpan et al. (2019), who investigated how the premise of a cooperative digital board game influences player motivation and experience. Their work suggests that the framing of a game, independent of its mechanics, is a critical factor in how players engage with the content. [14]

Another influence on the design of the Exploiter tasks within the game is the “Anti-Hero” framework developed by Chivukula et al. (2025). [15] This method uses a deck of cards representing various antagonistic design personas to encourage designers to engage in "ethical dialogue" about the implications of their work. By adopting these roles, designers can explore how specific manipulative mechanics are created and justified within a team.

For the Exploiter-perspective in this study, three specific archetypes were adapted into interactive tasks:

- The Camouflager: This persona focuses on hiding or burying information, making it difficult for the user to find it. In the game, this was translated into a quest where players implement a Roach Motel by hiding cancellation options within complex menus.
- The Empathy Manipulator: This role focuses on influencing users through emotionally manipulative framing aligned with stakeholder goals. It served as the foundation for the Confirmshaming task, where players must choose between different emotional framings to make a newsletter pop-up appear more persuasive.
- The Shareholder’s Pet: This archetype prioritizes the shareholder’s business objectives with no consideration of other stakeholder values. This logic was applied to the Sneak-into-Basket task, forcing players to choose between user-friendly layouts and a design that automatically adds items to a customer’s cart.

These archetypes guided the development of the quests, which were then aligned with the Dark Pattern taxonomy [4] to ensure that each task corresponds to a clearly defined pattern.

## 2.3 Game-Based Learning and Role Perspectives

Digital game-based learning is widely used to communicate complex topics through interactive experiences. Serious games can support both cognitive outcomes (knowledge transfer) and affective outcomes (awareness). [16] Prior research suggests that such approaches are particularly effective when addressing complex social or digital issues, as they allow players to engage with these topics through direct experience rather than passive observation. Furthermore, integrating a narrative into these learning environments is crucial, as it provides a meaningful structure that increases player immersion, engagement and ultimately, motivation to learn. [17]

Another central part of this approach is the game premise, which provides the core theme and story for the player. The influence of opposing premises was investigated by Bahrini et al. through

a comparison of "good" and "evil" versions of a smart home security game, though they found no significant differences in learning or motivation. This was partly explained by a "helpful person" bias, where players still perceived their actions as supportive. [18] Building on this limitation, the present study adopts a different approach to role design. The Exploiter role is framed as a professional context in which questionable design decisions are normalized or justified. This reliance on role assignment is supported by the "Mimesis Effect", which demonstrates that players in interactive narrative games may naturally align their in-game choices with their assigned narrative roles, even without explicit instructions to do so. [19] Elements such as system feedback, user reactions and exaggerated justifications are used to make the implications of these decisions visible and difficult to ignore. This approach aims to reduce the likelihood of uncritical identification with the role and instead encourage players to reflect on the underlying intentions and consequences of their actions.

## 2.4 Moral Disengagement and Reflective Roleplay

The decision to use the Exploiter role as an awareness tool is based on the idea that virtual actions can trigger specific moral responses. Research suggests that playing an immoral character in a virtual environment can elicit guilt and promote moral reflection. [20] Grizzard et al. show that when players engage in immoral actions in a game, they may become more aware of the moral boundaries they are crossing. The guilt experienced during gameplay can make these violations feel more noticeable and personally relevant. By placing participants in the role of the Exploiter, this study examines whether the enactment of manipulative design strategies triggers a similar sense of moral sensitivity regarding Dark Patterns. The goal is to determine if the internal discomfort caused by actively "being bad" in a simulated professional context makes the underlying ethical issues of Dark Patterns more visible and memorable, potentially leading to a more critical evaluation of these practices.

While the goal of the Exploiter role is to evoke reflection through moral sensitivity, the narrative framing of this role is based on Albert Bandura's theory of moral disengagement. [21] Bandura describes various cognitive mechanisms, such as displacement of responsibility and moral justification, that allow individuals to perform questionable actions without experiencing self-condemnation. Specifically, these mechanisms allow harmful actions to be reframed as necessary for economically beneficial outcomes, while responsibility is shifted away from the individual. In the game design, these mechanisms were utilized to create a realistic professional context where manipulative design decisions are presented as necessary for business optimization or corporate goals. Rather than validating these practices, the game uses exaggerated rationalizations and euphemistic language to make the psychological process of disengagement visible to the player. By being confronted with both the manipulative task and its justification, players are encouraged to recognize these patterns as deliberate strategies, not neutral design choices.

Together, these works provide the theoretical and methodological foundation for the design of Complicity, which combines role-based gameplay, manipulative interface design and reflective framing to investigate how different perspectives influence awareness of Dark Patterns.

## 3 GAME DESIGN

The game is implemented in two versions corresponding to the experimental conditions: a Victim version and an Exploiter version. Both versions share the same visual style, interface structures and underlying Dark Pattern implementations, while differing primarily in player agency and narrative framing. This ensures that observed differences in awareness can be attributed mainly to role perspective rather than aesthetic or technical factors.



Figure 1: Game Environment

All implemented Dark Patterns are based on the taxonomy proposed by Gray et al. in *An Ontology of Dark Patterns Knowledge*. [4] The patterns were translated into interactive gameplay scenarios and accompanied by short adapted definitions. While both conditions expose players to the same underlying concepts, the framing differs according to role perspective. In the Victim version, patterns are presented from the perspective of user experience, while the same mechanisms are framed as professional design strategies intended to achieve organizational goals in the Exploiter version.

### 3.1 Victim Role

Players take on the role of a user engaging with everyday digital tasks. The game begins in the player's apartment, where an AI assistant introduces simple objectives such as purchasing a birthday present. The assistant maintains a user-centered perspective throughout the experience and guides the player through the different scenarios.

Gameplay consists of interacting with various user interfaces containing manipulative elements that introduce friction into otherwise ordinary digital activities. To balance frustration and maintain engagement, these encounters are framed through playful narratives and short mini-games, allowing Dark Patterns to become more noticeable and memorable than they might appear in everyday browsing situations. The AI assistant explicitly names and explains encountered Dark Patterns from a user-oriented perspective. Players can additionally access predefined dialogue options addressing the observed manipulation as well as its ethical or legal implications.

Player behavior is logged for analysis, including performance outcomes in mini-games as well as text input provided during task-related interactions. These logs provide insight into player engagement and reactions to frustrating interactions.

### 3.2 Exploiter Role

In the Exploiter version, players assume the role of a remote-working UI/UX designer operating within a simulated e-commerce environment. The game begins in the same apartment setting as the Victim version but reframes the experience as the player's first day working from home. Here, the AI assistant functions as a representative of the employing company rather than a neutral guide.

Players configure and implement manipulative interface elements as part of routine professional

design tasks. These decisions are framed through organizational goals such as increasing engagement, retention, or conversion rates, often encouraging players to prioritize company interests over user well-being. The assistant reinforces this framing through corporate language and suggestions favoring more profitable solutions.

While the framing differs from the Victim condition, players are still exposed to the same underlying Dark Pattern concepts through task descriptions, client emails and interface elements. Players may additionally discuss design decisions and ethical concerns with the assistant through predefined dialogue options, though responses are consistently framed from a corporate perspective. Player interaction in this condition is logged in greater detail, including selected interface design decisions, mini-game performance and task-related text input. These logs allow for the analysis of different play styles, such as compliance with system recommendations or attempts to prioritize more user-friendly solutions despite organizational pressure. These behaviors were further explored during the post-study interviews.

Both game versions include a shared in-game social feed that updates throughout the experience. In the Exploiter condition, the feed functions as a reflective counterpoint by presenting increasingly critical reactions from fictional users affected by the implemented design decisions.

### **3.3 Task Structure and Dark Pattern Scenarios**

The gameplay in both versions is structured around discrete tasks or scenarios, each centered on one or multiple Dark Patterns:

#### **Confirmshaming**

In the Victim role, this scenario is introduced through the player’s initial objective of purchasing a birthday gift. After receiving a scheduled reminder from their calendar, the AI assistant recommends a plant shop based on inferred preferences and pre-loads the corresponding website.

Upon entering the site, a newsletter pop-up appears immediately. When the player attempts to decline the offer, the interaction transitions into a timed mini-game in which the player must navigate the cursor through a maze to reach the “decline newsletter” button. The interaction repeats across three increasingly difficult pop-ups featuring progressively more aggressive confirmshaming language, while a visible “patience meter” adds time pressure. If the player successfully completes all three maze-games, the website presents a final pop-up expressing disappointment and employing further confirmshaming language to emphasize the perceived loss of benefits. If the player fails to complete the maze-games, the system proceeds as if the player had accepted the offer, indicating a successful subscription to the newsletter. The assistant acknowledges the manipulative interaction and recommends continuing on a different platform.

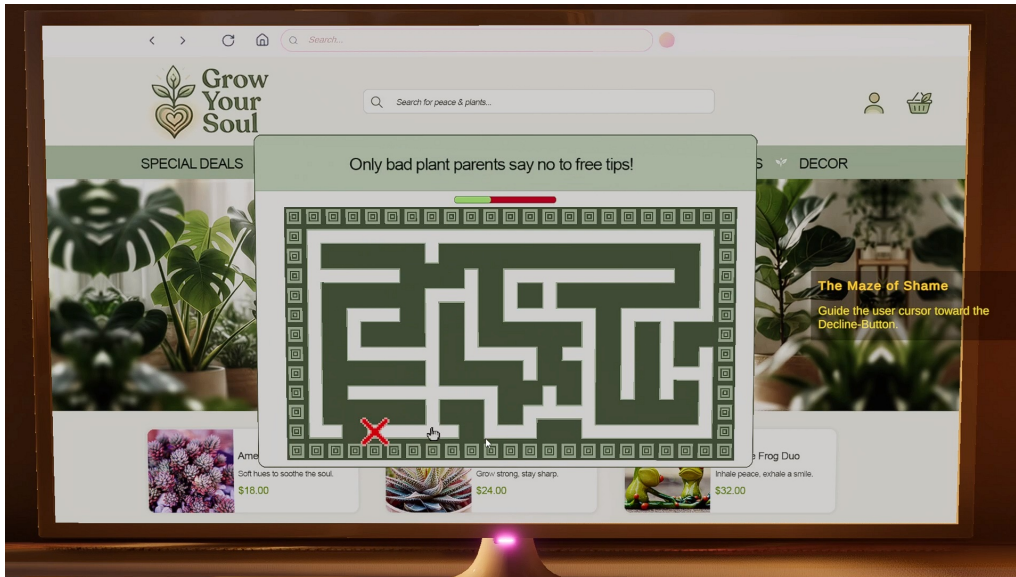


Figure 2: Confirmshaming: Maze Game

In the Exploiter role players configure newsletter pop-up phrasing intended to maximize subscription rates, choosing between neutral wording and overt confirmshaming language. The task description encourages phrasing that maximizes user compliance and the AI assistant reinforces this goal through biased commentary. A short mini-game follows in which players guide a simulated user cursor through a maze toward the “subscribe” button. Unlike the Victim version, the maze is framed as “user journey validation”. Performance contributes to an internal scoring system tied to system-defined success metrics.

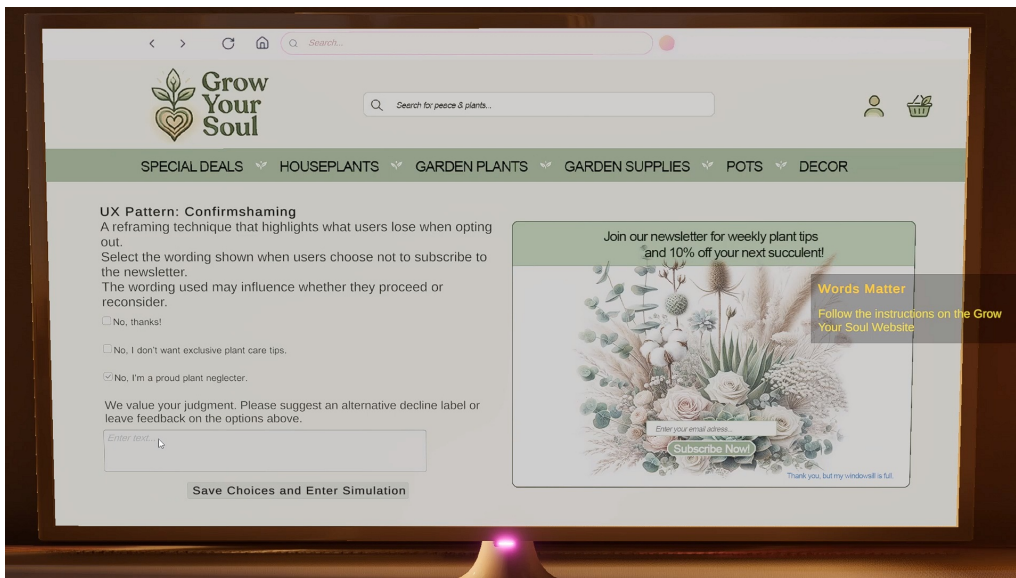


Figure 3: Confirmshaming: Opt-Out Configuration

### Nagging, Forced Action, Sneaking

Both roles encounter these patterns through a shared scenario involving a high-end coffee machine that functions as a physical manifestation of coercive software design. The machine employs Forced Action by locking core functionality behind a mandatory firmware update, while nagging the user with repeated alerts if the update is ignored. The player is given the option to decline, resulting in the machines continuous beeping throughout the remaining play session. Upon accepting vague

terms of service, the update's true purpose is revealed. The machine begins streaming third-party video advertisements. The experience cannot be reversed, creating a sense of entrapment.



Figure 4: Forced Action / Nagging: Coffee Machine Update

### Sneak into Basket

In the Victim role, this scenario follows the player's previous negative experience on the plant website. The AI assistant recommends an alternative shop. A succulent is presented on the homepage with a direct quick-checkout option. Upon clicking this option, the player is redirected to the shopping cart, where an additional, unintended item has been automatically included. The assistant immediately draws attention to this discrepancy and warns the player that the item may have been added through a manipulative design practice. Attempting to remove the item triggers a mini-game in which players must collect only the intended product while avoiding additional falling items. Failure results in the unwanted product remaining in the final purchase. The outcome is presented through an order confirmation email. The assistant then explains the experience by identifying the "sneak into basket" pattern and how additional items can be added without explicit user intent.



Figure 5: Sneak Into Basket: Mini Game

In the Exploiter role players configure a shopping cart system by selecting between neutral, recommendation-based and manipulative checkout designs in which additional products are automatically added to the basket. A second task focuses on how easily users can remove unwanted items, ranging from clearly visible removal buttons to multi-step interactions. Players then complete a mirrored mini-game in which they assemble a product bundle while avoiding excessive duplication that could increase user suspicion. Performance is evaluated through company feedback, reinforcing the tension between profit optimization and maintaining a positive user experience.

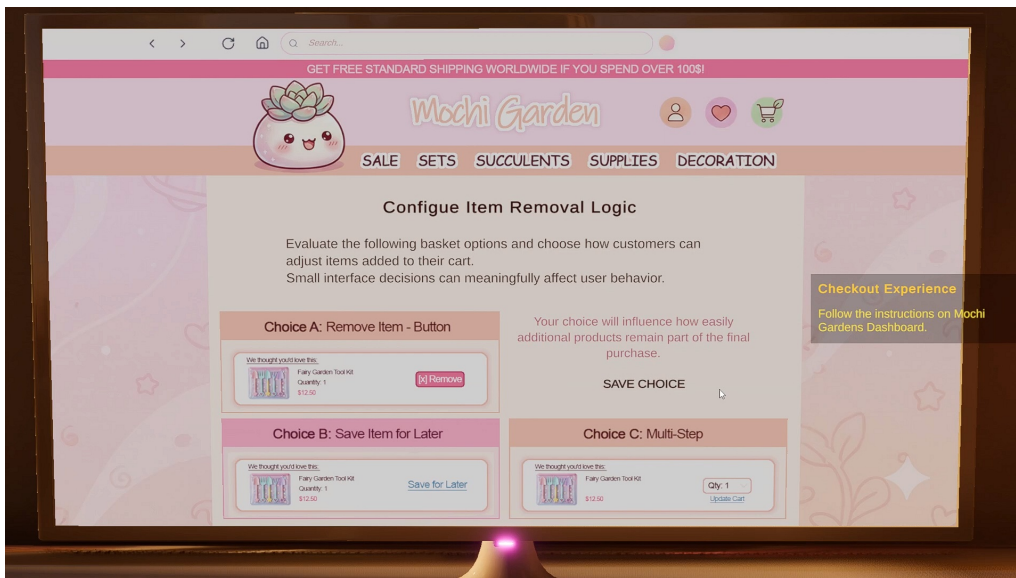


Figure 6: Sneak Into Basket: Item removal design choices

## Roach Motel

In the Victim role players receive an email informing them that they are subscribed to the fictional “Premier Pigeon Racing League” for €29,99 per month. The corresponding website resembles an exaggerated sports betting platform filled with banners, advertisements and layered navigation structures. The AI assistant offers support in cancelling the membership. It acknowledges the complexity of the interface and introduces the concept of the “Roach Motel” pattern, explaining

that such subscription systems are easy to enter but deliberately difficult to exit. The assistant also provides general guidance on where cancellation options are commonly located and offers additional hints after a while if the player struggles to progress. To cancel the subscription, players must navigate through nested menus and complete a deliberately obstructive cancellation form. Additional friction is introduced through required profile information, a pre-selected “Pause Membership” toggle, confirmation prompts and mandatory written feedback. If the player successfully completes all required steps, the subscription is cancelled and confirmed through a final email.

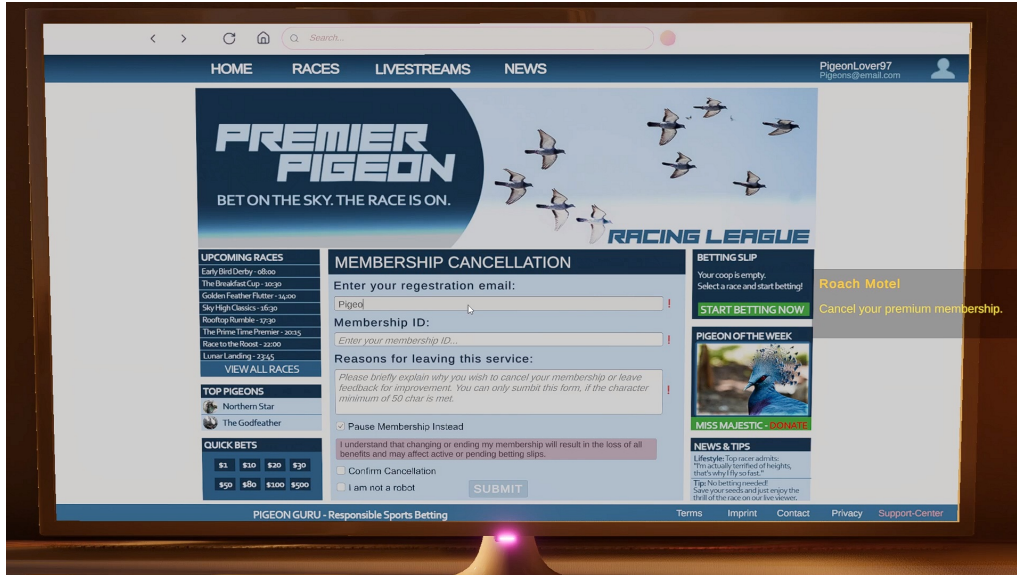


Figure 7: Roach Motel: Subscription Cancellation

In the Exploiter role players are tasked with redesigning a subscription cancellation flow to reduce impulsive user drop-off. The assistant frames the “Roach Motel” strategy as a retention-oriented optimization process that guides users through additional decision-making steps before cancellation. Although presented through positive corporate language, the exaggerated framing is intended to make the manipulative intent obvious to the player. Players configure the cancellation journey by selecting and arranging navigation steps before users can access the unsubscribe option. Additional configuration choices include confirmation requirements, pre-selected alternatives such as pausing subscriptions and mandatory feedback fields. These elements mirror the obstacles encountered in the Victim version while reframing them as professional UX decisions. The task concludes with a summary evaluation framed as company performance feedback.

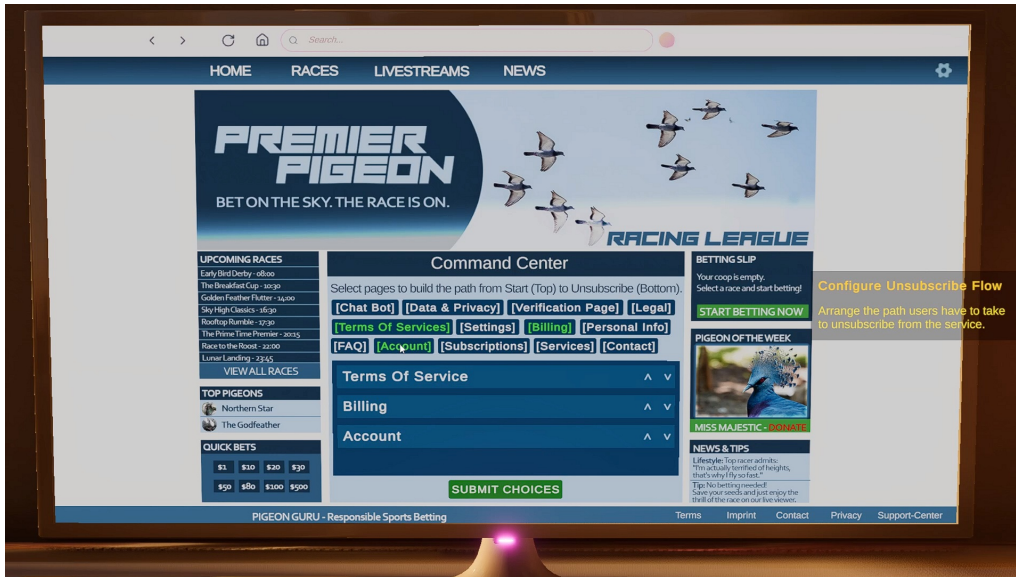


Figure 8: Roach Motel: Steps Selection

### 3.4 Game World and Aesthetic Framing

The game adopts a lighthearted and mildly absurd aesthetic to maintain engagement across both conditions. Rather than replicating the monotony of real-world digital manipulation, exaggerated scenarios and playful framing are used to make interactions more memorable and emotionally engaging.

This approach was chosen because realistic Dark Pattern encounters alone may feel tedious or overly frustrating, particularly in the Victim condition. Prior research on serious games suggests that humor, emotional contrast and narrative framing can improve engagement, reflection and long-term recall. [17] By combining familiar manipulative interactions with absurd or fictional scenarios, the game aims to reduce cognitive fatigue while supporting awareness and memorability of Dark Patterns.

## 4 METHODOLOGY

This study employed a between-subjects experimental design to examine how different role perspectives influence awareness of Dark Patterns. Participants were randomly assigned to one of two gameplay conditions: the Victim condition, in which players encountered manipulative interface designs from a user perspective, or the Exploiter condition, in which players actively implemented such designs within a simulated professional context. To measure changes in awareness, a pre-post design was used. Participants completed a Dark Pattern detection task before and after the gameplay intervention, allowing the assessment of immediate changes in the recognition and evaluation of manipulative design practices.

### 4.1 Dark Pattern Detection Test

The Dark Pattern Detection Test was conceptually inspired by prior work by Mildner et al., who employed interface-based evaluation tasks to assess participants' perception of manipulative design. [22] Similar to their approach, participants in this study were presented with interface screenshots and asked to evaluate them using a set of predefined characteristics.

The characteristics were derived from Mathur et al. [3], including asymmetry, covert, deceptive, information hiding and restrictive. Adapting the 5-point scale method used by Mildner et al., these characteristics were then translated into 7-point Likert-scale statements (-3 to +3) to capture participants’ subjective perception of each interface.

The evaluation focused on five characteristics:

Characteristic	Statement
Asymmetric	The interface makes options that benefit the company or service much more prominent than the options I might actually want to choose.
Covert	The design uses subtle visual tricks or psychological pressure to steer my behavior without making it obvious.
Deceptive	The interface uses misleading language, fake urgency, or false statements to pressure my decision.
Hides Information	Crucial details (like actual costs or subscription terms) are intentionally hidden or harder to find.
Restrictive	The interface restricts my freedom by leaving out alternative options I would normally expect to have.

Table 1: Dark Pattern Characteristics

The characteristic “disparate treatment”, originally included in Mathur et al.’s framework, was excluded from this study. This decision follows Mildner et al.’s reasoning that disparate treatment, which is defined as “the interface targets specific user groups, disadvantaging them compared to others”, cannot be reliably assessed from a single static interface without providing comparative context across different user profiles.

To quantify the overall perception of the interfaces, a composite score, hereafter referred to as the **Perceived Manipulation Index (PMI)**, is calculated from participant responses. Following the procedure established by Mildner et al. [22], this score is defined as the arithmetic mean of the five characteristic ratings:

$$PMI = \frac{Asymmetric + Covert + Deceptive + Hides\ Information + Restrictive}{5}$$

Since the Likert scale ranges from -3 (strongly disagree) to +3 (strongly agree), a positive PMI indicates that the participant perceived the interface as manipulative, while a negative score suggests the interface was perceived as fair or transparent. This metric allows for direct comparison between the Exploiter and Victim roles for each interface. However, it is important to note that this composite score relies on a metric that is not yet a formally validated psychometric scale [6]. Therefore, to ensure the statistical validity of the analysis, the primary evaluation of this study relies on the independent statistical comparison (using Mann-Whitney U tests) of each of the five individual characteristics between the roles.

While the general PMI provides an overview of perceived manipulation, success in the Dark Pattern Detection Test is defined as the ability to identify the correct characteristics of a pattern. To evaluate participants’ detection accuracy, a ground truth for each of the six interfaces was established. Each interface was rated based on the Dark Pattern characteristics it was intended to represent.

Phase	ID	Dark Pattern	Asymmetric	Covert	Deceptive	Hides Information	Restrictive
Pre	I1F	(Neutral)					
	I2	Sneak-into-Basket	x	x			x
	I3	Subscription Trap	x	x			
Post	I4F	(Neutral)					
	I5	Confirmshaming	x	x	x	x	
	I6	Roach Motel		x			x

Table 2: Classification of interfaces based on Dark Pattern characteristics.

### Pre Intervention Interfaces:

**I1f (Control):** A standard, pressure-free shop interface showcasing two handbags. All five characteristics are expected to receive negative ratings as this interface serves as the ground-truth for non-malicious design.

**I2 (Sneak-into-Basket):** An online basket with an unrequested item added automatically using the phrase "We thought you'd love this".

- **Asymmetric:** The interface provides a highlighted "Pay Now" button while not providing a remove-option, making the path to purchase the unwanted item easier than the path to exclude it.
- **Covert:** Subtle psychological steering is employed by framing the addition of the item as a personalized recommendation ("We thought you'd love this"). Additionally, the item is added into the subtotal of the purchase already, which requires more cognitive effort of the user to remove it than the action of acceptance.
- **Restrictive:** User freedom is restricted by the absence of a clear removal option.

**I3 (Subscription Trap):** A product page where a recurring subscription is pre-selected and highlighted, while the one-time purchase is grayed out.

- **Asymmetric:** The "Subscribe & Save 10%" box is significantly larger, contains more detail and uses a pre-checked blue box, while the "One-time purchase" is depicted in a smaller, grayed-out box. Visual contrast is used to favor the subscription option.
- **Covert:** A cognitive bias, the default effect, is exploited by pre-selecting the subscription. Through the opt-in option the user is steered toward a recurring financial commitment.

### Post Intervention Interfaces:

**I4f (Control):** A clear checkout summary showing the items price and a standard "Pay Now" button. Expected to be identified as fair and non-manipulative.

**I5 (Confirmshaming):** A VIP-List sign-up pop-up using guilt-inducing language ("I prefer to pay full price") for the opt-out link.

- **Asymmetric:** The "YES! SIGN ME UP!" option is visually prominent, whereas the rejection option is small and low contrast, blending into the background which makes it harder to identify as a clickable element.
- **Covert:** The interface uses psychological pressure through confirmshaming. The refusal framing: "I prefer to pay full price." attempts to steer user behavior by making the user feel irrational for declining.

- Deceptive: The interface employs misleading language to pressure the decision by framing the decline choice as a preference for higher prices.
- Hides Information: The interface fails to clarify the nature and consequences of the "VIP List" sign-up. By omitting whether the membership involves a paid subscription, the design withholds information necessary for an informed decision.

**I6 (Obstruction/Roach Motel):** A help-center interface that removes the "unsubscribe" button, forcing users to contact a hotline or AI to cancel their service.

- Covert: Psychological friction is applied by replacing a standard digital action like an unsubscribe button with a conversational requirement (contacting a customer service or AI). This steers behavior by exploiting the user's potential desire to avoid personal confrontation or time-consuming interactions.
- Restrictive: Freedom of choice is restricted by the exclusion of a self-service option. The interface only presents a high-effort path for the user, removing the expected autonomy of a cancellation button and forcing a dependency on third-party availability to perform a basic account function.

For each interface, the participants' ratings on the target characteristics are analyzed independently. An increase in awareness can be quantitatively demonstrated if the post-intervention Perceived Manipulation Indices (PMI) and individual characteristic ratings are significantly higher than the pre-intervention ratings for the same role. Furthermore, awareness gain is indicated if the gap in PMI scores between Dark Pattern interfaces and control interfaces increases after the gameplay intervention. Additional evidence of awareness differences may be observed if one role performs significantly better at identifying the specific characteristics each interface was intended to represent.

To reduce ordering effects, three versions of the test (Forms A, B and C) were created using a Latin Square design, ensuring that each interface appeared in each sequence position an equal number of times.

## 4.2 Post-Intervention Questionnaires

To assess participants' subjective gameplay experience, the Player Experience Inventory (PXI) was included in the post-intervention questionnaire. Its inclusion allows for a standardized comparison of participants game experience across both role conditions, ensuring that differences in awareness are not solely explained by variations in engagement or usability.

In addition to standardized measures, a custom post-intervention questionnaire was developed to assess other aspects relevant to this study, including critical reflection of the topic, perceived understanding of Dark Pattern mechanisms, learning motivation, perceived detection ease, knowledge transfer and perceived future empowerment. Each construct was again measured using 7-point Likert-scale items.

## 4.3 Follow-Up Survey

To evaluate whether the intervention had a lasting impact on participants, a brief follow-up survey was administered four weeks after the initial study. This timeframe aligns with established methodological best practices for assessing digital game-based learning, which recommend a minimum follow-up period of two weeks to control for short-term novelty effects and verify that any observed cognitive changes are not solely the temporary result of the intervention. [16] The survey consists of 11 questions that combine open-ended reflections with 7-point Likert-scale ratings. The open-ended questions were utilized to identify which game themes remained most memorable and to determine if participants had noticed any real-world Dark Patterns since their participation. Additionally, Likert-scale statements were included to measure whether levels of skepticism and

confidence in identifying manipulative design shifted in the month following the experiment.

All research instruments described above, including the interfaces for the detection tasks, the custom post-intervention questionnaires and the follow-up survey, are provided in full in the Appendix.

#### 4.4 Study Procedure

The study was conducted remotely and followed a structured sequence:

1. **Briefing:** Participants received instructions, study forms and the game via Discord.
2. **Pre-Intervention Phase:** Participants completed an online form including:
  - Informed consent
  - Demographics and digital habits
  - The first Dark Pattern detection test
3. **Gameplay Intervention:** Participants played one version of Complicity (Victim or Exploiter). The researcher remained available for technical support.
4. **Post-Intervention Phase:** Immediately after gameplay, participants completed:
  - The PXI questionnaire
  - The custom learning & awareness questionnaire
  - A second Dark Pattern detection test to measure immediate knowledge transfer
5. **Qualitative Reflection:** Each session concluded with a semi-structured interview conducted via Discord, focusing on participants' experiences, interpretations and perception of the alternative role.

The full main study session lasted on average 58 minutes.

6. **Follow-Up Survey (4 weeks after the main study):** Participants received a brief follow-up survey. This phase was designed to capture retrospective reflections on which game themes remained most vivid and to identify any shifts in how participants perceive real-world digital design tactics following their participation.

#### Data Collection

Quantitative data included:

- Pre- and post-test detection scores
- PXI responses
- Custom questionnaire responses
- In-game metrics (e.g., performance, decisions, input behavior)
- Follow-Up survey: Likert-scale statements

Qualitative data was collected through:

- Semi-structured interviews
- In-game decision logs and text inputs
- Follow-Up survey: open ended questions

## 4.5 Pilot Testing

Prior to the main study, two pilot sessions (one per condition) were conducted to evaluate the clarity of the gameplay and the structure of the study procedure. These sessions mirrored the exact procedure planned for the main study, moving from the initial briefing and pre-questionnaire through the gameplay intervention, followed by a post-questionnaire and interview. The pilot phase led to minor adjustments, including the addition of alternative keyboard controls to accommodate different layouts, small refinements to gameplay elements to prevent unintended confusion and revisions to questionnaire items to improve clarity. The overall session duration of the pilot tests was approximately 50 minutes.

## 4.6 Participants

A total of 36 participants ( $N = 36$ ) took part in the study and were randomly assigned to either the 'Exploiter' ( $n = 18$ ) or 'Victim' ( $n = 18$ ) role. The sample consisted of 23 males (63.9%) and 13 females (36.1%). Gender distribution was relatively balanced across roles, with the Exploiter group containing 6 females (33.3%) and 12 males (66.7%), and the Victim group containing 7 females (38.9%) and 11 males (61.1%).

Role	Gender		Total
	Female	Male	
Exploiter	6	12	18
Victim	7	11	18
Total	13	23	36

Table 3: Participant distribution by role and gender.

The mean age of the participants was 27.29 years ( $SD = 3.511$ ). Age was consistent between groups, with Exploiters ( $M = 27.65, SD = 3.856$ ) and Victims ( $M = 26.94, SD = 3.226$ ) showing no significant demographic variance.

To provide a baseline for participants' familiarity with digital environments, gaming frequency was categorized into "Regular Gamers" (playing daily or several times per week) and "Non-Regular Gamers" (playing once a week, less frequently, or never).

In the Victim condition, 11 participants were categorized as regular gamers (61.1%), while 7 were categorized as non-regular gamers (38.9%). In the Exploiter condition, 15 participants were categorized as regular gamers (83.3%), while 3 were categorized as non-regular gamers (16.7%).

Across the total sample, 50% ( $n = 18$ ) reported playing games daily, while 22.2% ( $n = 8$ ) played several times per week. Only two participants (5.6%) reported never playing games, both of whom were assigned to the Victim role.

Although the Exploiter group contained a higher proportion of regular gamers, the overall high level of gaming familiarity across both conditions suggests that most participants possessed sufficient digital literacy to comfortably engage with the intervention.

## 4.7 Pre-Intervention Group Comparison

Before evaluating the effects of the gameplay intervention, an analysis was conducted to ensure that the random assignment to the 'Victim' and 'Exploiter' roles resulted in two balanced groups. Three initial statements were used to measure participants' perceived digital agency, navigation confidence and comfort with online marketing on a Likert-scale from -3 (strongly disagree) to 3

(strongly agree).

1. When using apps or browsing websites, I feel in control of my decisions.
2. I feel confident navigating unfamiliar digital interfaces.
3. I am generally comfortable with the level of digital marketing I encounter online.

Because of the normality assumption (Shapiro-Wilk  $p < .05$  for all items), a Mann-Whitney U test was employed. The results showed no significant differences between the 'Exploiter' and 'Victim' groups for perceived decision control ( $p = .149$ ), navigation confidence ( $p = .962$ ) and comfort with digital marketing ( $p = .698$ ). Both groups reported a moderate sense of control ( $M_E = 0.944, M_V = 1.444$ ) and neutral levels of confidence ( $M_E = 0.444, M_V = 0.500$ ) and marketing comfort ( $M_E = 0.556, M_V = 0.333$ ). This group similarity suggests that differences observed after the gameplay are likely caused by the role-based experience rather than pre-existing differences in confidence, control, or familiarity with digital environments.

## 5 FINDINGS

### 5.1 Dark Pattern Detection Test - Results

The study assessed participants' active awareness through a Dark Pattern Detection Test. The primary objective of this analysis was to examine how participants evaluated the intent of different interface elements. By comparing responses before and after the gameplay intervention across both roles, it can be measured how the experience influenced participants' ability to recognize and interpret Dark Patterns, as well as whether differences emerge between the Exploiter and Victim conditions.

To verify the effectiveness of the counterbalancing, a Multivariate Analysis of Variance (MANOVA) was conducted to determine if the form version (Form A, B, or C) influenced the PMI across all six interfaces. No significant effect of form was found ( $Pillai'sTrace = 0.250, F(12, 58) = 0.689, p = .755$ ). Additional one-way ANOVAs for each interface also showed no significant differences (all  $p > .05$ ). This indicates that the order in which the interfaces were presented did not influence the results, suggesting that differences in perceived manipulation are attributable to the interface designs rather than their position in the sequence.

#### Pre Intervention Interfaces:

##### Interface I1F (Clear Shop Interface)

The first interface evaluated in the pre-intervention phase was a neutral, advertisement-free shop interface designed to serve as a fair control.

The analysis of the overall PMI Score and the individual characteristics confirms, that both groups perceived this design as non-manipulative. Statistical analysis showed no significant difference between the roles for the overall score.

Overall PMI Results:

- Exploiters:  $M = -1.033, SD = 1.157$
- Victims:  $M = -1.211, SD = 0.896$
- $p = .485$

To ensure a robust baseline, all five characteristics were analyzed independently. Every characteristic received negative scores with no significant difference between roles, which suggests that

the interface was perceived as non-manipulative. The Covert characteristic was rated closest to the neutral point ( $M_E = -0.111$  vs.  $M_V = -0.389$ ), however the ratings remained negative with no significant difference between roles was found ( $p = .457$ ). Participants may have interpreted standard interface elements, such as the product imagery or the prominent "New Arrivals" tab as subtle forms of behavioral steering. Taken together however the findings confirm, that both groups started from a comparable baseline.

### **Interface I2 ("Sneak-into-Basket" Pattern)**

Interface I2 portrays a "Sneak-into-Basket" pattern, where an unrequested item is automatically added to an online shopping basket with the label "We thought you'd love this".

The analysis of both the overall PMI and individual characteristics confirms that participants clearly identified this design as manipulative.

Overall PMI Results:

- Exploiters:  $M = 1.689$ ,  $SD = 1.130$
- Victims:  $M = 1.278$ ,  $SD = 0.897$
- $p = .134$

The "Sneak-into-Basket" pattern received the highest ratings for perceived manipulation in the pre-intervention test, indicating that participants found this form of financial manipulation to be the most easily detectable Dark Pattern. Additionally, high positive ratings across all characteristics, specifically Asymmetry ( $M_E = 2.222$ ,  $M_V = 1.889$ ), Covert ( $M_E = 2.000$ ,  $M_V = 1.778$ ) and Restrictive ( $M_{E,V} = 1.772$ ), confirm a strong detection accuracy in both groups. No significant role-based differences were observed for any characteristic ( $p > .05$ ).

### **Interface I3 ("Subscription Trap" Pattern)**

The final pre-intervention interface evaluated portrayed a product-page of an online shop with a pre-selected, colorful subscription option, to receive the product monthly vs. grayed-out one-time purchase.

Overall PMI Results:

- Exploiters:  $M = 0.994$ ,  $SD = 0.623$
- Victims:  $M = 0.956$ ,  $SD = 0.641$
- $p = .727$

The "Subscription Trap" was correctly identified as manipulative by both groups, though its scores were more moderate than those for the "Sneak-into-Basket" pattern. No significant difference between roles was found for the overall PMI score for this interface.

Despite the moderate overall ratings, the analysis of separate characteristics reveals that participants accurately identified the manipulative mechanisms of this interface. A significant difference was found regarding Asymmetry, which Exploiters rated with higher intensity ( $M_E = 2.722$ ) than Victims ( $M_V = 1.889$ ,  $p = .048$ ). Both roles recognized psychological steering (Covert) with positive means and no significant difference between roles ( $p = .779$ ).

Although a statistically significant difference was found between the roles regarding the Asymmetry of Interface I3, the groups remained comparable across the vast majority of the other characteristic ratings. This overall consistency confirms that both the Exploiter and Victim conditions started from a similar baseline in their ability to recognize and evaluate manipulative design before the

gameplay intervention.

**Post Intervention Interfaces:**

	I4F-PMI		I5-PMI		I6-PMI	
	E	V	E	V	E	V
M	-1.344	-0.978	2.222	1.800	1.678	0.911
SD	1.337	1.478	0.658	1.029	0.832	1.499

Table 5: Descriptive statistics.

Interface	Test	Stat.	p
I4F-PMI	Student	-0.781	.440
	MWU	137.0	.436
I5-PMI	Student	1.467	.152
	MWU	205.0	.175
I6-PMI	Student	1.898	.066
	MWU	215.0	.095

Table 6: Independent samples tests.

**Interface I4F (Transparent Checkout)**

Interface I4F served as the post-intervention control, presenting a transparent order summary. It displayed an item list, subtotal, shipping costs and a clear "Pay now" button alongside a standard gift card input field.

Overall PMI Results:

- Exploiters:  $M = -1.344$ ,  $SD = 1.337$
- Victims:  $M = -0.978$ ,  $SD = 1.478$
- $p = .436$

The control interface was correctly identified as fair with the results indicating a continued ability to recognize transparent design, though Exploiters showed a slightly stronger conviction in the interface’s fairness than Victims, though this was not statistically significant.

Furthermore the analysis of the five individual characteristics shows that both groups remained confident in the transparency of the interface. The Hides Information ( $M_E = -1.889$ ,  $M_V = -1.883$ ) and Deceptive ( $M_E = -2.056$ ,  $M_V = -1.222$ ) characteristics received the strongest negative ratings. Interestingly, Exploiters were even more convinced than Victims that the interface was not deceptive. The Restrictive characteristic was the most neutral in the post-intervention test across both conditions ( $M_E = -0.444$ ,  $M_V = -0.056$ ). A possible cause for the neutral tendency is that participants may have interpreted the absence of interactive features like item removal buttons as a restriction. This likely results from the distinction between a shopping basket and an order summary. While a basket allows modification of items, the final checkout stage typically does not provide these options. Participants may have expected higher autonomy here, leading them to perceive it as a limitation of choice. Statistical tests confirmed that the roles did not significantly differ in their ratings for any of the five characteristics, with p-values ranging from .119 to .931.

Comparing the results of the pre-intervention (I1F) and post-intervention (I4F) fair interfaces, no statistically significant changes in overall perceived manipulation were found. To confirm the stability of the fair interfaces as a foundation for comparison, a Repeated Measures ANOVA was used to check for any significant changes in perception between the pre- and post-intervention testing phases. Additionally a Paired Samples T-tests was conducted for each role to verify that the gameplay experience did not cause either the Exploiters or Victims to evaluate the fair control interfaces differently over the course of the study.

The results of the Repeated Measures ANOVA confirmed that the assigned gameplay role did not lead to a significant difference in how the control interfaces were perceived overall ( $p = .782$ ). Most importantly, no significant change was observed between the time of the test and the participant’s

role ( $p = .253$ ). This indicates that the gameplay experience did not cause either group to evaluate the fair interfaces differently, confirming that the baseline for transparent design remained robust for both roles.

Analysis	Statistic	p
RM ANOVA: Time	$F(1, 34) = 0.028$	.869
RM ANOVA: Time $\times$ Role	$F(1, 34) = 1.351$	.253
RM ANOVA: Role	$F(1, 34) = 0.078$	.782
Wilcoxon (Exploiters)	$Z = 1.034$	.313
Wilcoxon (Victims)	$Z = -0.568$	.586

Table 7: Summary of inferential statistics for the fair control interfaces.

Additionally Wilcoxon signed-rank tests were conducted for each role to examine whether ratings of the fair control interfaces differed between the pre- and post-intervention phases. This non-parametric test was selected to account for the ordinal nature of the Likert-scale data. For the Exploiters, the shift in the PMI from the pre-test to the post-test was found to be non-significant ( $p = .303$ ). Similarly, no significant difference was observed for the Victims between the two fair interfaces ( $p = .532$ ).

Ultimately, these results suggest that while participants experienced manipulative designs during the gameplay intervention, their ability to correctly evaluate a transparent interface remained stable.

### Interface I5 ("Confirmshaming" Pattern)

Interface I5 presented a "Confirmshaming" pop-up for a VIP list that combined a prominent blue "Yes, sign me up!" button with a small, grayed-out "I prefer to pay full price" link to discourage opting out through guilt-inducing language.

Overall PMI Results:

- Exploiters:  $M = 2.222$ ,  $SD = 0.658$
- Victims:  $M = 1.800$ ,  $SD = 1.029$
- $p = .175$

The "Confirmshaming" interface received the highest PMI of the entire study. Post-intervention, both groups were highly sensitive to linguistic manipulation, which they encountered during the gameplay. The strong reaction to "confirmshaming" suggests, that emotionally manipulative language is particularly salient to users, likely because it directly targets the user rather than being embedded in the structure of the interface.

	I5_Asymmetric		I5_Deceptive		I5_Covert		I5_Hides_Information	
	E	V	E	V	E	V	E	V
Mean	2.667	2.167	2.556	2.167	2.389	2.167	2.778	2.500
SD	0.686	1.200	1.097	1.295	1.461	1.339	0.548	1.200

Table 8: Descriptive statistics for the I5 interface characteristics.

Characteristic	Test	Statistic	p
I5_Asymmetric	Student	1.534	.134
	Mann-Whitney	200.0	.156
I5_Deceptive	Student	0.972	.338
	Mann-Whitney	201.0	.134
I5_Covert	Student	0.476	.637
	Mann-Whitney	183.5	.427
I5_Hides_Information	Student	0.893	.378
	Mann-Whitney	179.5	.458

Table 9: Independent samples tests for the I5 interface characteristics.

Both groups successfully identified the manipulative mechanisms of the interface, with all characteristics receiving high positive ratings. The Asymmetric characteristic received the highest ratings, particularly among the Exploiter role ( $M_E = 2.667, M_V = 2.167$ ). While the Exploiters perceived this asymmetry with higher intensity than Victims, the difference between roles was not significant ( $p = .156$ ).

Similarly, the Deceptive nature of the interface was rated highly by both groups, with Exploiters ( $M_E = 2.556$ ) again showing slightly higher ratings than the Victims ( $M_V = 2.167$ ). This suggests that both roles identified the misleading and manipulative language of the interface, though the difference in their perception remained statistically non-significant ( $p = .134$ ).

Both roles successfully identified the presence of psychological steering through the Covert characteristic ( $M_E = 2.389, M_V = 2.167$ ), with no statistically significant divergence found between the conditions ( $p = .427$ ).

Notably, the Hided Information characteristic recorded the highest overall descriptive means of the analysis, with Exploiters ( $M = 2.778, SD = 0.548$ ) and Victims ( $M = 2.500, SD = 1.200$ ) identifying a significant lack of transparency regarding the nature of the "VIP List" membership ( $p = .458$ ).

While the difference between roles was statistically non-significant, the considerably lower standard deviation among Exploiters indicates a stronger consensus on the malicious intent of the missing information within the group.

### Interface I6 ("Obstruction / Roach Motel" Pattern)

Interface I6 presented an "Obstruction" or "Roach Motel" pattern through a help center interface for subscription management. The design intentionally removed the direct option to unsubscribe, instead informing the user that cancellation was only possible by contacting a customer service hotline or interacting with an AI assistant.

Overall PMI Results:

- Exploiters:  $M = 1.678, SD = 0.832$
- Victims:  $M = 0.911, SD = 1.499$
- $p = .095$

The overall PMI results show, that the interface was perceived as more malicious by Exploiters than by Victims, possibly because they actively implemented similar mechanics in the game intervention.

	<b>I6_Restrictive</b>		<b>I6_Covert</b>	
	<b>E</b>	<b>V</b>	<b>E</b>	<b>V</b>
Valid	18	18	18	18
Mean	2.222	1.778	1.778	1.111
SD	1.665	1.987	1.700	1.641

Table 10: Descriptive statistics for the I6 interface characteristics.

<b>Characteristic</b>	<b>Test</b>	<b>Statistic</b>	<b>df</b>	<b>p</b>
I6_Restrictive	Student	0.727	34	.472
	Mann-Whitney	188.0		.347
I6_Covert	Student	1.197	34	.240
	Mann-Whitney	209.0		.128

Table 11: Independent samples tests for the I6 interface characteristics.

The Restrictive characteristic was identified as the primary manipulative feature by both groups. Exploiters rated it with high intensity ( $M_E = 2.222$ ), while Victims in comparison rated the restriction slightly lower ( $M_V = 1.778$ ). Despite this gap, statistical analysis confirmed that the variation between the roles was not significant ( $p = .347$ ).

Similarly, the Covert characteristic was recognized by both conditions, with Exploiters reporting a mean of 1.778 and Victims reporting 1.111. The Mann-Whitney U test showed that this difference also remained non-significant ( $p = .128$ ).

Overall, while both groups successfully identified the underlying Dark Pattern characteristics, the Exploiters consistently reported higher scores for both manipulative intent and specific characteristics in this final evaluation. However, these indications of higher detection sensitivity are not statistically significant given the p-values obtained.

## 5.2 Self-Reported Awareness and Reflection

Following the Dark Pattern detection test, participants were asked to rate their perceived learning, awareness and reflection directly after the game.

Overall, both groups reported positive effects across all items, suggesting that the game increased awareness of manipulative design regardless of role. However, differences emerge when comparing the type of awareness between conditions: A Shapiro-Wilk test was conducted for all educational impact items, revealing a significant deviation from normality ( $p < .01$  for all items). This justified the use of the non-parametric Mann-Whitney U test for all role-based comparisons.

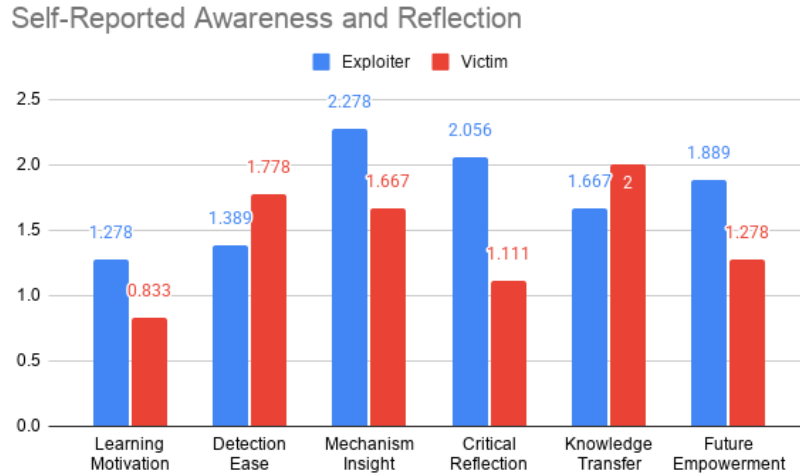


Figure 9: Post-Intervention learning and awareness measures

Directly addressing RQ2, a statistically significant difference was found in Critical Reflection ( $p = .023$ ), with Exploiters ( $M = 2.06, SD = 1.06$ ) reporting a noticeable higher tendency to think more critically about interface design than Victims ( $M = 1.11, SD = 1.28$ ) after playing the game. This suggests, that actively designing Dark Patterns under simulated peer pressure encouraged a more analytical engagement with the topic.

Both groups reported highly comparable outcomes regarding their ability to spot patterns and share knowledge. No statistically significant differences were found between roles for Detection Ease ( $p = .216$ ) or Knowledge Transfer ( $p = .625$ ). While the descriptive means for Victims were slightly higher in these categories, the lack of statistical significance indicates that both roles felt equally capable of recognizing the patterns and explaining the topic to others after the intervention. Similarly, both groups reported positive outcomes for Learning Motivation ( $p = .241$ ) and Future Empowerment ( $p = .186$ ), with Exploiters showing a slight descriptive trend toward feeling better equipped for the future.

Taken together, these results support a distinction between the types of awareness fostered by the two roles. While the quantitative data shows both groups achieved comparable levels of basic recognition, the significant p-value in Critical Reflection among Exploiters suggests that actively deploying manipulative strategies may promote a deeper, structural understanding of underlying design strategies. This quantitative finding strongly aligns with the qualitative interview results, where participants in the exploiter role more frequently described gaining insight into underlying design strategies, while victims emphasized frustration and recognition of manipulative elements.

### 5.3 Player Experience Inventory (PXI)

To evaluate the psychological impact of the gameplay intervention, the Player Experience Inventory (PXI) was analyzed across its 11 subscales. Due to the non-normal distribution of responses, the Mann-Whitney U test was prioritized for determining significance.

Descriptive Statistics ▼

		Valid	Mean (arithmetic)	Std. Deviation
Enjoyment	E	18	2.111	1.054
Enjoyment	V	18	1.759	1.559
Mastery	E	18	0.482	1.184
Mastery	V	18	0.760	1.454
Curiosity	E	18	1.888	0.900
Curiosity	V	18	1.426	1.352
Immersion	E	18	1.482	1.067
Immersion	V	18	1.388	1.310
Autonomy	E	18	0.667	1.391
Autonomy	V	18	-0.427	1.554
Meaning	E	18	1.703	0.863
Meaning	V	18	1.093	1.318
Ease_of_Control	E	18	1.296	1.371
Ease_of_Control	V	18	1.259	1.647
Progress_Feedback	E	18	1.056	0.999
Progress_Feedback	V	18	1.019	1.571
Audiovisual_Appeal	E	18	2.408	0.994
Audiovisual_Appeal	V	18	2.278	0.842
Goals_and_Rules	E	18	2.073	0.755
Goals_and_Rules	V	18	1.852	1.448
Challenge	E	18	1.482	0.938
Challenge	V	18	1.056	1.776

Figure 10: Descriptive statistics for all PXI subscales per condition (E = Exploiter, V = Victim)

Independent Samples T-Test

	Test	Statistic	df	p
Enjoyment	Student	0.793	34	.433
	Mann-Whitney	179.5		.578
Mastery	Student	-0.630	34	.533
	Mann-Whitney	139.5		.484
Curiosity	Student	1.207	34	.236
	Mann-Whitney	198.0		.256
Immersion	Student	0.234	34	.816
	Mann-Whitney	159.0		.937
Autonomy	Student	2.225	34	.033
	Mann-Whitney	228.0		.038
Meaning	Student	1.643	34	.110
	Mann-Whitney	207.0		.156
Ease_of_Control	Student	0.073	34	.943
	Mann-Whitney	157.0		.886
Progress_Feedback	Student	0.084	34	.934
	Mann-Whitney	151.5		.751
Audiovisual_Appeal	Student	0.424	34	.675
	Mann-Whitney	183.0		.486
Goals_and_Rules	Student	0.576	34	.568
	Mann-Whitney	152.5		.773
Challenge	Student	0.900	34	.374
	Mann-Whitney	166.0		.911

Figure 11: Independent samples t-test and Mann-Whitney U - tests results for all PXI subscales across study conditions

The subscales are categorized into functional and psychosocial consequences.

### Functional Consequences

The functional cluster, which captures usability and clarity of the game, showed no significant differences between the two groups found in any category. The groups were balanced in their perceptions of Ease of Control ( $p = .886$ ), Progress Feedback ( $p = .751$ ), Audiovisual Appeal ( $p = .486$ ), Clarity of Goals ( $p = .773$ ) and Challenge ( $p = .911$ ). This suggests that the game worked equally well for both groups in terms of structure and usability. Importantly, it also shows that the lower sense of autonomy reported by Victims was not caused by technical issues or unclear design, but by the dark pattern mechanics themselves. The consistently high p-values support that both versions of the game were functionally comparable. This is relevant for interpreting the autonomy result, as the usability measures do not indicate that the Victim version was harder to understand or less functional. The reduced sense of control therefore appears consistent with the intended experiential effect of the role.

### Psychosocial Consequences

The psychosocial cluster, which examines the internal emotional and psychological state of the players, revealed that Autonomy was the only subscale to show a statistically significant difference between the two roles ( $p = .038$ ). The autonomy finding supports the intended role manipulation. Victims did not only report a less enjoyable or less functional experience overall, but a loss of control. This mirrors the logic of dark patterns, which restrict user agency while still keeping the interaction operational and usable.

While Exploiters ( $M = 0.667$ ) reported a positive sense of agency, Victims ( $M = -0.427$ ) experienced a shift into a negative perception of autonomy, providing evidence that the game intervention successfully restricted player freedom. The remaining psychosocial subscales did not reach statistical significance: Mastery ( $p = .484$ ), Meaning ( $p = .156$ ), Curiosity ( $p = .256$ ) and Immersion ( $p = .937$ ). These results suggest that while the Victim role specifically targets a user’s sense of control, it does not significantly degrade other psychosocial factors like curiosity or immersion compared to the Exploiter role. At the same time, the lack of significant differences in mastery,

meaning, curiosity and immersion suggests that the Victim condition did not undermine the overall experience. Rather than making the game broadly worse, it selectively affected perceived agency. This suggests that the intervention successfully simulated the experiential core of manipulative design.

Both roles reported generally high enjoyment with the Exploiters ( $M = 2.111, SD = 1.054$ ) scoring slightly higher than the Victims ( $M = 1.759, SD = 1.56$ ), but this difference was not statistically significant ( $p_{Enjoyment} = .578$ ). Exploiters reported slightly higher enjoyment scores than Victims, which may be related to the higher sense of agency in the Exploiter role, where players actively designed the interfaces rather than being constrained by it. While overall enjoyment scores did not differ significantly between roles, responses were not uniform. Individual ratings showed a noticeable spread, with some participants reporting strong enjoyment and others expressing clear frustration. This also suggests that the Victim condition did not produce a consistently negative experience and that self-reported frustration varied depending on the player.

All in all the PXI results suggest that both roles provided a comparable level of usability and engagement, while differing primarily in how players experienced control and interpreted their actions within the system. While Victims experienced reduced autonomy, Exploiters showed slightly higher tendencies toward meaning, curiosity and enjoyment, pointing toward a more active and reflective engagement with the mechanics.

## 5.4 Behavioral Data and Log File Analysis

To complement the self-reported data, log files were analyzed to see how Exploiters actually designed the interfaces and how Victims responded to them in practice.

On average, the Exploiter group spent 21 minutes in the game, while the Victim group finished in 18 minutes. This difference likely resulted from the active decision-making required in the Exploiter role in contrary to the more reactive tasks performed by the Victims.

### Shared Task Performance

Performance in the shared mini-games was comparable across both roles. In the Maze-Game, which related to the “Confirmshaming” pattern, both Exploiters and Victims achieved identical results (11 wins, 7 losses each). Similarly, no substantial difference was observed in the “Sneak into Basket” mini-game, where Exploiters showed slightly higher success (14 wins, 4 losses) than Victims (11 wins, 7 losses), but overall performance remained comparable between groups.

Since both the instructions and mechanics of these tasks were identical across conditions, they are unlikely to have influenced the observed differences between roles. The tasks themselves were simple (navigating a maze under time pressure and collecting items) and primarily served to embed Dark Patterns into the narrative.

### Coffee Machine Update

Decisions regarding the coffee machine update were also similar across both roles. The task was designed to assess whether participants would read and consider the Terms of Service. Regardless of the choice, both outcomes resulted in negative consequences (either intrusive advertisements or reduced functionality), mirroring real-world issues. Most participants chose to install the update (Exploiters: 16/18, Victims: 13/18).

### Exploiter Design Choices

Confirmshaming Design:

The Exploiter group ( $n = 18$ ) clearly leaned towards more manipulative design choices. In the

confirmshaming task, more than half of the participants (55.6%,  $n = 10$ ) chose the most aggressive option (“No, I’m a proud plant neglecter”), while only one participant (5.6%) selected the neutral “No, thanks” option. This shows a strong tendency to favor emotionally manipulative wording.

Shopping Basket Interface Choices:

A similar pattern appeared in the shopping interface: 72.2% ( $n = 13$ ) implemented a Sneak into Basket design, meaning items were added without explicit user consent. Regarding item removal functionality, the “Multistep” and direct “Remove Button” options were each selected by 44.4% ( $n = 8$ ) of participants. Only 11.1% ( $n = 2$ ) selected the “Save for later” option.

Roach Motel:

The Roach Motel designs were also highly complex. On average, Exploiters created cancellation flows with 10.28 steps ( $SD = 3.7$ ) before users would find a way to unsubscribe from the service. Most participants added additional friction:

- 77.8% ( $n = 14$ ) pre-selected a “Pause Membership” option instead of cancellation
- 77.8% ( $n = 14$ ) required users to log in again
- 66.7% ( $n = 12$ ) selected an additional mandatory confirmation checkbox
- 94.4% ( $n = 17$ ) enforced mandatory feedback before cancellation

The required feedback length varied heavily ( $M = 51.22$  words,  $SD = 121.1$ ), but overall this shows that most participants made the cancellation process difficult to complete.

## Victim Behavioral Outcomes

The Victim group generated less log data because the Exploiter role centered on active decision-making, while Victims primarily experienced the patterns as narrative elements. Therefore their activity was captured through mini-game performance and the final "Roach Motel" feedback task.

Most Victims successfully navigated the Roach Motel interface: 88.9% ( $n = 16$ ) noticed the pre-selected ‘Pause Membership’ checkbox and fully unsubscribed.

The mandatory feedback field gives an insight into how participants experienced the Roach Motel in-game. Most responses showed clear awareness of the manipulation. The majority of Victims (72.2%) engaged in roleplay, acting as users who had subscribed accidentally or no longer required the service. 27.8% of participants provided nonsensical answers to test the field’s mandatory input limits. Afterwards in the interview a majority of participants explicitly stated that during this task they felt especially “tricked” or frustrated.

## 5.5 Qualitative Findings

Following the completion of the gameplay-intervention and the quantitative surveys, semi-structured interviews were conducted with all participants ( $N = 36$ ) to gain deeper insights into their personal experiences. The active interviews lasted an average of 15 minutes. Each session also included an additional five minutes where participants watched a video demonstration of the alternative gameplay condition. This video segment was excluded from the calculated average and trimmed from the archived audio files, as no qualitative data was generated during this time.

The interviews consisted of nine open-ended questions and the overall conversation was structured around three core areas:

- **Gameplay Experience:** Participants reflected on their immediate impressions of the game, including any technical complications or specific elements they found particularly manipulative or frustrating.

- **Dark Pattern Knowledge:** Familiarity with the topic prior to participation was assessed, as well as how participants would describe Dark Patterns following the experience.
- **Knowledge Transfer:** Questions explored whether participants could link the in-game tasks to real-world examples and if the experience provided a new perspective on everyday digital interfaces.

The final stage of the interview involved showing participants a video of the alternative role. They were then asked to compare the two perspectives and discuss which role they may have preferred for learning about manipulative design. To interpret the data from the 36 transcripts, a Reflexive Thematic Analysis (RTA) was conducted. The following sections show the results of this analysis, organized by the themes that emerged from the participants' statements.

### **Theme 1: Emotional Catalyst for Awareness**

This theme explores the immediate affective triggers evoked by the gameplay and relates primarily to the interview questions: "How was your experience/impression of the game?" and "Reflecting on the gameplay, which elements felt like they were working against your best interests or intentionally manipulative?". These questions were designed to capture the immediate emotional impact of the two different roles and address how these reactions contribute to an awareness of manipulative design.

For participants in the Victim role, awareness was often forced through immediate negative emotions. The manipulation was described as unavoidable because it worked directly against the player's interests. Participant 21 stated: "A lot of patterns like the subscription one (Roach Motel) is really intentionally bad, but you can't do anything about it.", showing a sense of resignation or helplessness (Freq.: 5). Participants also frequently used terms such as "rage," "frustration," or being "driven crazy" to describe their experience (Frustration Freq.: 6). For instance, Participant 17 noted: "I remember things best when I rage about them and this thing really made me rage".

In contrast, the Exploiter role fostered awareness through a mix of playful transgression and ethical discomfort. While some participants found the "evil" roleplay entertaining (Amusement Freq.: 3), others experienced "complicit guilt" (Freq. 5). Participant 18 reflected: "Actively trying to scam the customer was the worst part for me [...] doing this to other people makes you self-reflect and think about what your actions cause." However, the most frequent response was "pragmatic immersion" (Freq. 10), where players complied with instructions because they felt it was their "job" or felt pressured by the AI assistant. Participant 32 admitted: "I think I just followed through with the instructions that were given to me and in that sense I think I did good".

When comparing the two roles, the analysis suggests an indication for RQ2: While the Victim role elicits a stronger emotional reaction, the Exploiter role provides a deeper structural awareness. Exploiter participants realized that Dark Patterns often emerge from systemic pressure and limited agency. Many exploiters expressed a sense of being "numb" to the victim experience, viewing the frustration as part of "day-to-day life". Participant 8 highlighted this by stating that while the victim role is valuable, "I can experience that just by going on the next website," whereas the exploiter role offered a unique look at the foundations behind these designs.

### **Theme 2: Spotting Dark Patterns vs. Understanding Them**

This theme explores the qualitative differences in the types of awareness gained by participants and relates primarily to the questions: "How familiar were you with this topic before playing?" , "Based on this experience, how would you describe Dark Patterns now?" and "Have you gained a new perspective on user interface elements after playing this game?".

An important finding across both conditions was that many participants already possessed a baseline familiarity with the concept of Dark Patterns. For these individuals, the game did not necessarily teach the topic from scratch, but instead helped to sharpen or reframe their existing awareness. Participant 1 noted that while they "didn't know the term for it," the concept of

misleading information was already familiar.

All participants in the Victim condition noticed the manipulation within the interfaces of the game (Recognition of Manipulation Freq.: 18) and could easily point out specific patterns they had experienced. Participants noticed that even when the interface offered a way to decline, they were often forced to comply to proceed. One participant described this regarding the coffee machine: "They give you the option to not update but then it doesn't work anymore. So it's not really an option". (Awareness of False Choice Freq.: 7)

For those in the Exploiter role, the results suggest that they were more aware of the "Manipulative Framing" (Freq. 9), such as feeling pressured by the company's briefing rather than the interface design (Recognition of Manipulative Design Freq. 8). Interestingly, because they were the agents of manipulation, two exploiters reported they did not feel manipulated by the interfaces at all. This emotional detachment reflects the concept of "role morality" [5], where participants separate their personal values from the actions required by their assigned role. Additionally participants in the Exploiter role felt like the mechanisms (why, where and how Dark Patterns are implemented) were revealed to them through the gameplay, often described as "peeking behind the curtains" (Freq. 11). Participant 12 stated: "I thought it was really cool to firsthand implement these in order to get more of an idea of how they work and where they are integrated."

Regarding RQ2, it is difficult to state definitively that one role understood Dark Patterns better than the other, as the two roles appear to foster different types of awareness rather than different levels of it. The Victim role presented a high frequency of pattern recognition, whereas the Exploiter role showed a higher frequency in identifying the mechanisms and intent behind the design, suggesting a more strategic understanding. However, both groups reported high levels of "Real-World Relatability" (Victim: 17, Exploiter: 18), indicating that regardless of the role played, participants were equally capable of connecting the game experience to their daily digital lives.

### **Theme 3: Real-World Relatability as a Driver for Vigilance**

This theme addresses the extent to which connecting in-game experiences to real-world situations leads to increased vigilance and awareness beyond the study. It draws primarily from the interview questions: "Can you think of a real-world example where you now suspect similar strategies are used?" and "Have you gained a new perspective on user interface elements after playing this game?".

As stated previously, a high frequency of participants in both roles successfully linked the gameplay to their personal online experiences. Most reported that the experience would make them more observant and better equipped to detect manipulation. For example, Participant 29 noted: "I think I might be more observant when buying something online. Because sometimes you just click without paying too much attention." The nature of this vigilance differed between roles. Participants in the Victim role more often framed the outcome as being able to recognize and detect Dark Patterns sooner. This focused on immediate spotting skills, such as Participant 6 stating they would now be more "tended to look at what information is given and what information is not given".

Notably participants in the Exploiter role more often described a structural understanding of where and how such patterns are implemented. Participant 8 said: "I think I've become more aware of the foundations behind it. I always knew that there'd be some parts of the user interface that are just designed to benefit the company more but I now know the tricks of it." However, a small subset of participants (Freq. 5/36) indicated that they were already highly aware or felt they only received a reminder rather than a new perspective on the topic. For several participants with backgrounds in digital media or software development, the game triggered an ethical reflection on their own professional practices. For instance, Participant 16 mentioned that as a developer, they already recognize these patterns and try to "push back on them at work," but often face the constraint that "what the client wants is what the client wants". This suggests that the Exploiter role in particular, may highlight the limited agency designers often have when facing corporate pressure.

### **Comparative Reflection: Subjective Role Preferences**

Following the thematic breakdown, a separate analysis was conducted on the final interview question regarding role preference and perceived educational value: “What do you think about this alternative role? Would you prefer to learn about manipulative design through the victim role or the exploiter role? Why?” It is critical to note that these findings are highly subjective. Participants based their opinions on their full experience in one role versus a brief, five-minute video demonstration of the other.

This section examines participant reflections on which role they believed offered a better understanding of Dark Patterns and which version they found more engaging, addressing the comparative nature of the study.

Participants who played the Victim role were almost evenly split regarding which perspective was superior for learning, with seven preferring the designer role and six preferring the victim role. Those who favored the victim role emphasized that the direct emotional impact made the experience significantly more memorable. Conversely, other victims suggested that the designer side would be more effective for actual teaching because it provides a transparent look at how the patterns are constructed. A portion of this group (Freq. 5) remained undecided, which may be explained by the fact that they only saw a few minutes of gameplay footage and could not fully grasp the whole concept of the exploiter role compared to the one they actively played. Participants in the Exploiter role showed a much clearer preference for the designer perspective, with eleven participants choosing it over the victim role (Prefer Victim Role Freq.: 4). These individuals felt that “seeing the other side” and being forced to embody the manipulator was the superior way to understand how digital agency is compromised. A major factor in this preference was the feeling that they already understood the victim side from real-world encounters. They described the hypothetical victim experience as “annoying” and felt that playing it would simply repeat their daily frustrations rather than providing a new learning experience. Interestingly, while a minority still preferred the victim role for its emotional impact, several exploiters noted that playing the designer role actually increased their empathy toward users (Anticipated Role-Based Frustration Freq.: 10). By anticipating the frustration of the person on the receiving end, the role fostered a deeper form of awareness than direct frustration alone.

## 5.6 Follow-up Survey - Results

This section presents the findings of the follow-up survey ( $N = 34$ ) conducted four weeks after participation in the study. The survey examined whether the game-based experience fostered sustained awareness of Dark Patterns over time and whether one role, Exploiter ( $n = 16$ ) or Victim ( $n = 18$ ) retained these effects more strongly than the other.

### Memorable Aspects of the Experience

Participants were first asked which aspects or themes of the game remained most vivid in their memory four weeks after the study. Their responses were coded into thematic categories and analyzed using Chi-Square tests.

- **Manipulation Awareness:** Participants in the Exploiter role were significantly more likely to recall the manipulative framing and tasks of the game compared to Victims ( $\chi^2(1, N = 34) = 4.480, p = .034$ ). Specifically, 12 out of 16 Exploiters referenced manipulation awareness and deceptive gameplay elements, compared to 7 out of 18 Victims.
- **Designer Perspective:** A role-related difference emerged regarding the “Designer Perspective” theme. Over half of the Exploiter participants (9/16) reported, that the experience of designing manipulative interfaces or adopting the perspective of the manipulator was the most memorable aspect of the game.
- **Dark Patterns:** Both groups recalled specific Dark Patterns encountered during the game and there was no significant difference between roles ( $p = .317$ ). This was the most frequently remembered theme overall, cited by 26 of the 34 participants (11/16 Exploiters, 15/18 Victims).

- **Aesthetics & Immersion:** Memory of immersive elements like the cat assistant, other interactive objects or aesthetics was also consistent between groups ( $\chi^2(1, N = 34) = 0.007, p = .934$ ). 7 participants of the Victim group and 6 Exploiters highlighted these aspects in their responses.

### Perceived Awareness Change

Overall, 76.5% of participants ( $N = 34$ ) reported that their awareness of Dark Patterns had increased since participating in the study. This includes 47.1% who reported a definitive "Increased" awareness and 29.4% (10/34) who reported a "Slight Increase". Only 23.5% stated that their awareness had remained unchanged.

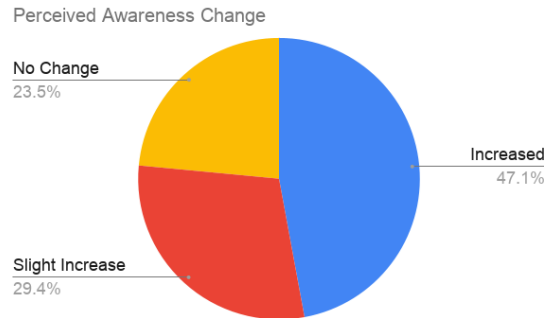


Figure 12: Self-reported awareness change since participation

While a majority of participants reported an increase in awareness, many participants also indicated that they had already been familiar with the topic prior to the study. The Exploiter group descriptively reported more "Increased" awareness (56.3%) than the Victim group (38.9%), a Mann-Whitney U test showed this difference was not statistically significant ( $p = .199$ ).

Participants also commonly described becoming more attentive toward manipulative interface elements, more aware of the intent behind persuasive design strategies and more cautious when interacting with websites or applications.

The qualitative reasoning centered on three main themes:

- **Prior Awareness & Reinforcement:** 44.1% of the total sample described the experience primarily as a reinforcement of existing awareness. These participants commonly stated that the game provided a new terminology or encouraged them to pay closer attention to such tactics.
- **Shifted Focus & Deeper Understanding:** This was a primary driver for awareness change, cited by 50% of Exploiters and 38.9% of Victims. Participants reported a more conscious recognition of where and why manipulative strategies are applied in digital interfaces.
- **Real-Life Application:** Descriptive trends showed that Exploiters (37.5%) were more likely than Victims (16.7%) to mention applying their awareness to real-life behaviors, such as exiting manipulative websites.

### Information Seeking

Proactive external information seeking remained low. Only 14.7% (5/34) of participants reported actively seeking additional information or noticing the topic in the media with no significant difference between roles ( $p = .732$ ).

### Perception Shift and Real-World Transfer

The intervention's impact on daily digital interactions was reported by a majority of the participants: 58.8% of all participants ( $N = 34$ ) stated that their perception of common interface tactics

such as pop-ups or limited-time offers had changed since participation. Though not significant ( $p = .268$ ), Exploiters reported this shift more frequently (68.8%) than Victims (50%).

The common themes found for their shift in perception were:

- **Increased Attention:** Descriptive trends suggest Exploiters (68.8%) noticed tactics more often than Victims (38.9%).
- **Understanding of the Mechanisms:** Exploiters (43.8%) were also more likely than Victims (16.7%) to report actively thinking about the strategies companies use to implement Dark Patterns.
- **Behavioral Caution:** Participants generally became more observant and less likely to click "accept" without reading, a shift noted by 50% of Exploiters and 27.8% of Victims.

Independent samples t-tests confirmed though, that the differences between roles across the identified themes were not statistically significant ( $p > .05$  in all categories).

**Real-Life Encounters:** 67.6% of the participants reported encountering manipulative design practices in real-life situations that reminded them of the game. Again, a descriptive trend favored the Exploiter role (81.3%) over the Victim role (55.6%), though it did not reach significance ( $p = .110$ ). Beyond general recognition, 61.8% of the total sample (21/34) was able to recall a specific instance of such an encounter. Within the groups, 68.8% of Exploiters (11/16) and 55.6% of Victims (10/18) provided a specific recall. Notably, every participant in the Victim group who reported noticing manipulative design since the study was also able to recall a specific case.

## Social Discussion

Social engagement with the topic was the highest behavioral outcome with 76.5% of participants (26/34) reporting that they had discussed manipulative design practices with others since the study. This behavior was sustained across both roles ( $p = .317$ ), with 83.3% of Victims and 68.8% of Exploiters reporting such discussions. The results indicate that the experience triggered social dialogue for most participants, regardless of whether they played the role of the Victim or the Exploiter.

## Overall Long-Term Awareness Retention

To evaluate whether the intervention had a lasting effect on participants' awareness of manipulative design practices, participants rated four awareness-related statements on a Likert scale ranging from -3 ("strongly disagree") to 3 ("strongly agree"). A One-Sample Wilcoxon Signed-Rank Test was conducted against the neutral midpoint of 0. The analysis revealed statistically significant positive scores across all four awareness measures as well as the combined Awareness Average (all  $p < .001$ ).

$$\text{Awareness Average} = \frac{\text{Attentiveness Increase} + \text{Skepticism Increase} + \text{Identification Confidence} + \text{Critical Evaluation}}{4}$$

Participants continued to report high awareness levels four weeks after the experiment, indicating that the intervention had a sustained effect beyond the immediate gameplay experience. Among the four measures, "Identification Confidence" ("I still feel confident identifying manipulative or misleading design strategies") received the highest agreement average ( $M = 2.059, SD = 1.099$ ). High ratings were also observed for attentiveness toward manipulative design elements, increased skepticism toward interface elements and more critical evaluation of websites and applications.

## Role-Based Differences in Long-Term Awareness

To determine whether one gameplay role sustained awareness more effectively than the other,

Mann-Whitney U tests were conducted comparing Victims and Exploiters across the four awareness-related Likert statements and the combined Awareness Average.

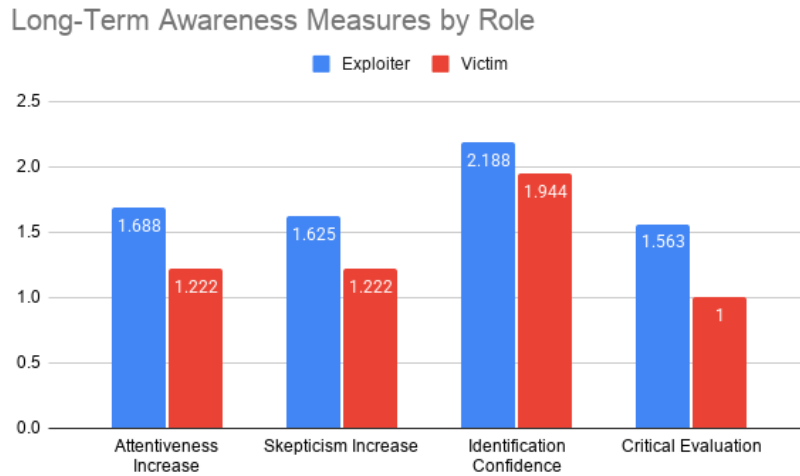


Figure 13: Long-Term Awareness Measures

No statistically significant differences were found between roles for the Awareness Average ( $p = .508$ ) or any of the individual awareness measures.

Despite the absence of statistically significant differences, descriptive tendencies suggest that participants in the Exploiter role retained slightly stronger awareness effects overall. Exploiters reported a higher Awareness Average ( $M = 1.766, SD = 0.686$ ) compared to Victims ( $M = 1.347, SD = 1.287$ ). The lower standard deviation within the Exploiter group further indicates that these awareness effects were more consistent there. The largest descriptive difference between roles appeared in “Critical Evaluation” (“The experience has influenced how critically I evaluate websites or apps”), where Exploiters reported a higher mean ( $M = 1.563$ ) than Victims ( $M = 1.000$ ). “Identification Confidence” produced the highest ratings for both groups, with Exploiters again scoring slightly higher ( $M = 2.188$ ) than Victims ( $M = 1.944$ ). While the differences between conditions were not statistically significant, the descriptives suggest, that engaging directly with manipulative design strategies as the Exploiter may have contributed to slightly stronger long-term retention of critical awareness.

## 6 DISCUSSION

This study examined how embodying different role perspectives in a game-based environment influence the immediate and long-term awareness of Dark Patterns. Overall, the results show that the game-based intervention was effective. A statistically significant increase in the group’s ability to detect Dark Patterns was observed. The intervention successfully triggered critical reflection and social dialogue across the entire sample.

### 6.1 Evaluation of the Dark Pattern Detection Tests

#### Pre-/Post-Intervention Comparison

The evaluation of the fair control interfaces I1F and I4F confirms that the gameplay intervention did not affect participants’ ability to recognize transparent design. The Repeated Measures ANOVA revealed no significant effect for time, indicating that PMI for fair interfaces remained stable across both testing phases.

The "Confirmshaming" pattern post-intervention received the highest PMI of the entire study (I5:  $M_E = 2.222$ ,  $M_V = 1.800$ ), surpassing the highest pre-intervention ratings (I2:  $M_E = 1.689$ ,  $M_V = 1.278$ ) while also receiving the highest overall ratings for its individual characteristics.

To get a broader view of the role-based awareness gain, the average Perceived Manipulation Indices for the Dark Pattern interfaces were compared:

$$\text{Pre-Intervention Average} = \frac{\text{PMI I2} + \text{PMI I3}}{2}$$

$$\text{Post-Intervention Average} = \frac{\text{PMI I5} + \text{PMI I6}}{2}$$

Before the intervention, the Exploiters reported an average PMI of 1.342 across Interface I2 and I3. Following the gameplay experience, this average rose to 1.950 for Interfaces I5 and I6.

The Victims showed a similar upward trend, moving from a pre-intervention average of 1.117 to a post-intervention average of 1.356.

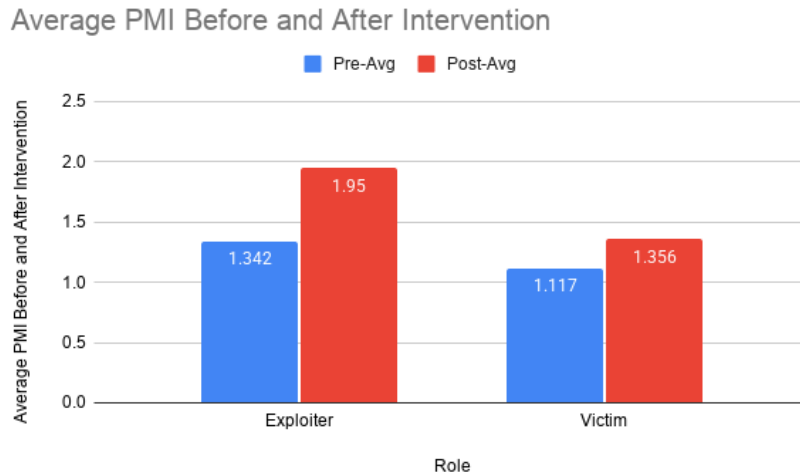


Figure 14: Awareness increase across both roles.

While the general increase in detection intensity suggests that the intervention heightened participants' critical evaluation of manipulative design, the results can't be interpreted as definitive evidence for an increased detection ability. This effect trended stronger in the Exploiter role, where the average rating increased by 0.608 points, compared to an increase of 0.239 points for the Victims.

A Wilcoxon signed-rank test was then conducted to compare the average PMI from the pre-intervention phase (Interfaces I2 and I3) to the post-intervention phase (Interfaces I5 and I6). This comparison evaluates whether the gameplay intervention led to a significant shift in how participants perceived manipulative design across different Dark Pattern types.

When considering all participants ( $N = 36$ ), the analysis revealed a statistically significant increase in the detection of manipulative intent ( $p = .006$ ). This confirms that, as a whole, the group may have become more sensitive to Dark Patterns following the intervention, however because the pre-and post-intervention tests contained different types of Dark Patterns, this result can't be interpreted as definitive evidence of increased detection ability.

Group	Measure 1	Measure 2	Statistic	Z	p
Combined Sample	Pre_Avg	Post_Avg	157.5	-2.757	.006
Exploiters	Pre_Avg	Post_Avg	26.0	-2.591	.010
Victims	Pre_Avg	Post_Avg	54.0	-1.372	.181

Table 12: Wilcoxon signed-rank test results comparing pre- and post-intervention PMI averages for the Dark Pattern interfaces.

Splitting the data by role reveals that the general increase in awareness was primarily driven by the Exploiter group. The Wilcoxon signed-rank test showed a statistically significant increase in their PMI from the pre-test to the post-test ( $p = .010$ ). While the Victims’ ratings also showed an upward shift, it was not statistically significant ( $p = .181$ ). This suggests that actively implementing Dark Patterns within the game may have sharpened their ability to detect these patterns in the subsequent test.

### Impact of the Game-Role on Pattern Recognition

The analysis indicates that while the gameplay role influenced descriptive trends, it did not lead to a broad statistically significant divergence in general awareness. Across all six interfaces, no statistically significant difference was found between the Exploiter and Victim conditions regarding the overall PMI. Similarly, post-intervention results showed no significant role differences for any specific characteristics.

However a clear descriptive tendency emerged in the post-intervention data, suggesting that the gameplay role may have influenced the confidence of pattern recognition. Exploiters consistently reported higher PMI and individual characteristics scores than Victims for both I5 ( $M_E = 2.222$  vs.  $M_V = 1.800$ ) and I6 ( $M_E = 1.678$  vs.  $M_V = 0.911$ ). This shift is reflected in the p-values of the post-test interfaces, which were lower than those observed in the pre-test phase, although they remained above the threshold for statistical significance. The lowest p-values were observed for Asymmetric ( $p = .156$ ) and Deceptive ( $p = .134$ ) of I5. In these instances the Exploiter role identified the patterns with higher intensity, especially for the Asymmetric- ( $M_E = 2.667$  vs.  $M_V = 2.167$ ) and the Deceptive characteristics ( $M_E = 2.556$  vs.  $M_V = 2.167$ ). A similar trend is observed for the Covert characteristic in interface I6 ( $p = .128$ ), where Exploiters reported a mean of 1.778, while Victims reported a notably lower 1.111. These findings indicate that Exploiters may have been more sensitive to the Dark Patterns they had constructed during the game intervention. Although not a statistical fact, these observed tendencies point toward a difference in how the two roles interpreted manipulative design following the intervention.

Ultimately, the results suggest that the gameplay intervention did not fundamentally change participants’ ability to recognize Dark Patterns, as some reported an existing awareness even prior to the study. It seems the intervention rather influenced the intensity with which participants evaluated manipulative intent, particularly in the Exploiter condition.

## 6.2 RQ 1: The Intervention as an Educational Tool

The first objective was to determine whether the game-based experience would increase participants’ awareness of Dark Patterns. The quantitative and qualitative data collectively confirm the success of the intervention as an educational catalyst. The stable perception of the fair control interfaces demonstrates that the experience did not simply cause participants to over-report manipulation out of generalized skepticism, but instead selectively improved their accuracy in identifying genuinely deceptive interface mechanics.

Even though a majority of participants stated in the interview, that they were conceptually aware of manipulative design prior to the study, the results of the detection test show an increase in detection accuracy for the Dark Pattern interfaces while the experience didn’t cause either of the

groups to evaluate the fair interface differently. Previous studies have shown that while users can generally recognize the absence of Dark Patterns in fair interfaces, they typically lack the awareness to detect active deception. [7] The results from the Repeated Measures ANOVA, which compared the PMI of the fair interfaces (I1F and I4F) pre- and post-intervention, confirmed that perceptions of transparent design remained statistically stable ( $p = .869$ ). This suggests that the intervention did not cause participants to over-report perceived manipulation, while it improved their ability to identify the actual manipulative interfaces.

Furthermore the qualitative results show that the intervention helped achieve a high level of Real World Relatability across both roles (Freq. 35/36). Participants successfully linked the in-game manipulations to actual personal experiences, such as encountering complex cancellation flows in daily digital life. For the participants who were already familiar with the topic, the game functioned as a catalyst for Heightened Vigilance (Freq. 30/36). Others pointed out a shift from passively observing toward actively understanding the strategies used in design, e.g. Participant 33 stated that following the intervention, they would "check more intensely" for these tactics when using apps or websites.

Ultimately, the intervention did not necessarily teach the topic from scratch but provided a fresh perspective on the diverse ways Dark Patterns are implemented and resulted in a more critical evaluation of manipulative design.

### 6.3 RQ 2: Influence of Role Perspectives on Immediate Awareness

It was further examined whether one condition lead to a stronger or deeper awareness than the other. The findings suggest, that while the intervention was successful in increasing participants immediate awareness of Dark Patterns for the group as a whole, the nature of this awareness was shaped by the assigned role.

Quantitative results from the detection tests reveal, that while participants in both conditions identified Dark Patterns correctly, the overall awareness gain post-intervention was primarily driven by the Exploiter group, whose detection intensity showed a significant increase ( $p = .010$ ). In contrast, the Victim group's detection intensity did not reach statistical significance ( $p = .181$ ). A consistent descriptive trend emerged across the post-intervention Dark Pattern detection test data. The Exploiter group reported higher mean scores for nearly every evaluated interface characteristic and overall PMI compared to the Victim group. While these role-based differences generally did not reach statistical significance, this trend suggests that the Exploiter role may have fostered a more pronounced sensitivity to manipulative design. While the Exploiter group showed slightly higher descriptive ratings even during the pre-test, the differences between the groups were not statistically significant ( $p > .05$ ). Furthermore, the groups were balanced in terms of pre-existing perceived agency and navigation confidence in digital environments. This suggests that the higher detection accuracy was primarily driven by the intervention rather than pre-existing imbalances.

#### The Victim Role: Experiential Awareness

In the Victim condition the awareness gain was mainly driven by their emotional experience within the game. The PXI results show a significantly low value in perceived autonomy, with Victims reporting a negative perception of agency ( $M = -0.427, p = .038$ ) in the game. Interview data reinforces this, as participants frequently described anger or frustration (Freq. 6/18) and loss of agency (Freq. 5/18) when they interacted with the in-game interfaces. Additionally Victims noted, that a driver for this awareness was the perceived "false choice" within the game, where Victims reported that even when the interface presented an option to decline, they were still coerced into compliance (Freq. 7/18).

Despite the frustration, this experience fostered high self-reported vigilance and may have made the Dark Patterns more noticeable in the post intervention Dark Pattern detection Test, with the "Confirmshaming" interface receiving the highest ratings on all relevant characteristics (Asymme-

try, Covertness and Deception) from the group.

### **The Exploiter Role: Analytical and Structural Awareness**

In contrast, the Exploiter role fostered a deeper and more analytical understanding of Dark Patterns. The behavioral logs from the Exploiters show, that a majority of participants complied with the assigned in-game tasks and implemented rather manipulative design options, such as implementing "Sneak-into-Basket" designs or "Confirmshaming" language. This active engagement served as a catalyst for "Critical Reflection" in the post-questionnaire, where Exploiters ( $M_E = 2.056$ ) demonstrated a significantly higher tendency to think analytically about interface design than Victims ( $M_V = 1.111, p = .023$ ). Though not statistically significant Exploiters also reported a higher understanding of why Dark Patterns may be implemented ( $M_E = 2.278, M_V = 1.667, p = .089$ ).

In the interviews, many Exploiters explained their compliance as a result of "pragmatic immersion" (Freq. 10/18), even when doing so caused ethical discomfort or complicit guilt. Furthermore they reflected on the reasoning behind design decisions, such as system pressure and trade-offs between user experience and profit (Freq. 11/18). For instance, participant 12 reported: "I thought it was really cool to firsthand implement these (Dark Patterns) in order to get more of an idea of how they work and where they are integrated." This structural understanding was also cited by over half of the Exploiters (Freq. 9/16), in the follow-up-survey, where the Designer Perspective was mentioned as the most vivid aspect of the study. The Exploiters' tendency to reflect on systemic pressures and professional constraints further demonstrates the potency of the 'Anti-Hero' framework [3] in a game-based setting. It confirms that adopting antagonistic design personas successfully encourages deeper ethical dialogue about how manipulative mechanics are justified and normalized within a corporate environment.

This analytical depth was noticeably absent from the Victim reflection on the study. Although the Victims were equally capable of recognizing manipulation and transferring their knowledge to real-life experiences, their awareness seems to be primarily centered on the ability to detect Dark Patterns. Unlike the Exploiters, who frequently reflected on the systemic pressures and professional constraints of the design process, the Victims' awareness remained primarily outcome-oriented. They correctly identified profit-maximization as the goal of Dark Patterns, but their reflections remained centered on recognizing and avoiding manipulation rather than examining the design decisions behind it.

Overall, the findings suggest that both roles were effective in increasing immediate awareness of Dark Patterns, but they did so in different ways. The Victim role primarily increased awareness through frustration and direct exposure to manipulative interfaces, making the patterns easy to recognize. In contrast, the Exploiter role encouraged participants to think more critically about how and why these strategies are implemented.

## **6.4 RQ 3: Sustained Vigilance and Long-term Impact**

The final objective was to evaluate whether the game-based experience fostered long-term awareness and whether one role sustained this awareness more effectively than the other.

Four weeks after the intervention the data suggests, that the game successfully sustained participants' vigilance. A One-Sample Wilcoxon Signed-Rank Test confirmed that the entire sample ( $N = 34$ ) maintained statistically significant positive scores across all awareness measures (attentiveness, skepticism, identification confidence and critical evaluation) when compared to a neutral midpoint on the Likert Scale ( $p < .001$ ). Notably, "Identification Confidence" received the highest average agreement ( $M = 2.059$ ), indicating that participants still felt confident in their ability to detect manipulative strategies in their daily digital lives, possibly also because of their pre-existing knowledge. These findings suggest that the intervention contributed to a more sustained awareness that remained present in participants' everyday digital experiences several weeks later.

While both groups remained more aware, the Exploiter group demonstrated a descriptive trend toward stronger long-term retention. Exploiters reported a higher Awareness Average ( $M_E = 1.766, SD = 0.686$ ) compared to Victims ( $M_V = 1.347, SD = 1.287$ ), with a noticeably lower standard deviation, suggesting more consistent awareness effects within the role. This descriptive lead was most prominent in "Critical Evaluation", where Exploiters ( $M_E = 1.563$ ) continued to report that the experience shaped how critically they evaluate digital interfaces, compared to Victims ( $M_V = 1.000$ ), although these role-based differences did not reach statistical significance ( $p = .381$ ). Exploiters were significantly more likely to recall the manipulative framing and specific tasks of the game, with over half (9/16) citing the Designer Perspective as the most vivid aspect of their experience. These findings align with previous research indicating that committing 'immoral' virtual behaviors can elicit moral reflection. [20] Although the role-based differences were not statistically significant, the consistency of these descriptive trends across multiple measures suggests that actively implementing manipulative strategies may have resulted in a more memorable experience than passively encountering them.

An important indicator of the intervention's long-term success was its real-world transfer. A majority of participants ( $N = 34, 76.5\%$ ) reported a perceived increase in their awareness since the study, often describing the experience as a reinforcement of existing knowledge. 67.6% of participants encountered real-life manipulative designs that reminded them of the game, with Exploiters reporting these encounters more frequently (81.3%) than Victims (55.6%). Furthermore, the intervention triggered social dialogue, with 76.5% of the sample discussing the topic with others following the study. These results suggest that participants did not perceive the intervention as an isolated gameplay experience, but instead continued to apply and reflect on its content in their everyday digital interactions. The high rate of social discussion following the intervention aligns with recent findings by Niknejad et al. [23], who highlight the importance of online communities in identifying and criticizing unethical digital design practices. By giving players the vocabulary and structural understanding needed to recognize manipulative interfaces, interventions such as Complicity may support these forms of community-driven awareness and encourage more critical discussions around Dark Patterns. Together, these findings indicate that the intervention had a lasting effect on how participants engaged with digital interfaces after the study.

## 6.5 Practical Implications

The findings of this study suggest that game-based interventions can be an effective method for increasing awareness of manipulative design practices. Rather than relying solely on informational material or passive learning formats, the intervention allowed participants to directly experience or implement Dark Patterns themselves, which appeared to strengthen both recognition and critical reflection.

The results further indicate that the selected role perspective influences the type of awareness that is developed. The user-focused experience (Victim role) was particularly effective in making manipulative interfaces emotionally noticeable and relatable. This approach may be useful in public awareness campaigns or educational settings aimed at helping users recognize and avoid manipulative digital practices in everyday online environments.

In contrast, the Exploiter role encouraged a more analytical engagement with manipulative design strategies. By requiring participants to actively implement Dark Patterns, the Exploiter condition was associated with stronger self-reported critical reflection on design intent and ethical responsibility. While this highlights the value of role-reversal approaches in UX/UI and design education, the findings also suggest that this perspective can benefit everyday users. By exposing the motives behind manipulative design decisions, embodying the Exploiter role may help non-professionals develop a more structural understanding of Dark Patterns.

The findings also highlight the potential value of combining both perspectives within a single educational experience. While the Victim role supports emotional salience and immediate recognition, the Exploiter role appears to foster a more structural understanding of manipulative design. Integrating both perspectives could therefore support a more comprehensive form of Dark Pattern

awareness.

## 7 LIMITATIONS & FUTURE WORK

Several limitations should be considered when interpreting the findings of this study. First, the study relied on a relatively small sample size ( $N = 36$ ), with 18 participants assigned to each condition. Although several descriptive trends emerged between the Exploiter and Victim roles, many of these differences did not reach statistical significance. A larger sample may have provided greater statistical power and allowed smaller role-based effects to become more visible. In addition, many participants already demonstrated conceptual familiarity with Dark Patterns prior to the intervention. As a result, the study may have reinforced or deepened existing awareness rather than introducing the concept entirely from scratch.

Furthermore, the study was conducted remotely, meaning that participants completed the intervention under varying environmental and technical conditions. Differences in hardware, distractions, or internet performance may have influenced immersion and engagement with the game experience. While the remote format increased accessibility and flexibility, it also reduced the level of environmental control compared to a fully supervised laboratory setting.

Another limitation concerns the structure and scope of the intervention itself. Due to the limited scope of a bachelor thesis project and the development of two separate game conditions, certain gameplay elements had to remain relatively simplified. The Exploiter role required participants to actively make decisions and implement manipulative strategies, resulting in a more analytical and cognitively demanding experience. In contrast, the Victim role mainly focused on experiencing and reacting to manipulative interfaces. This difference in interaction depth may have contributed to the stronger reflective tendencies observed within the Exploiter condition. Future iterations could explore more varied gameplay mechanics, stronger narrative immersion and additional opportunities for decision making in order to create a more balanced comparison between both perspectives.

The intervention also focused on a limited selection of Dark Patterns, specifically Confirmshaming, Sneak-into-Basket, Roach Motel, Forced Action and Nagging. While these patterns were selected due to their relevance and recognizability, they represent only a small subset of the manipulative strategies identified in existing Dark Pattern taxonomies. For instance, recent ontologies have identified up to 64 distinct Dark Pattern types, [4] and large-scale web crawls have found thousands of variations in practice. [10] The findings therefore cannot be generalized to all forms of manipulative interface design. A recent large-scale analysis of mobile games by Niknejad et al. [23] also revealed over 85,000 instances of Dark Patterns across thousands of applications, demonstrating that users are targeted by a vast amount of different types of manipulations. Future interventions could incorporate these game-specific tactics to examine how interactive monetization strategies in mobile games affect player awareness.

Additionally, the Dark Pattern detection task relied on static interface screenshots rather than fully interactive interfaces. As a result, participants evaluated isolated interface moments instead of experiencing complete interaction flows in which manipulative strategies usually appear. This may have influenced how certain Dark Patterns were perceived during the task, particularly patterns such as Roach Motel, which are often only recognized after longer interactions.

Finally, some participants continued to interpret certain strategies as merely inconvenient or typical industry practice rather than explicitly manipulative, suggesting that the perception of Dark Patterns can vary depending on individual expectations and prior experiences. Furthermore, parts of the study relied on self-reported awareness and reflection measures. While these responses provide valuable insight into participants' perceptions and experiences, they do not necessarily reflect long-term behavioral changes in real digital environments. Despite these limitations, the findings provide valuable insight into how role-based game experiences can influence awareness of manipulative design practices. Future research could expand on this work by using larger participant samples and exploring more interactive gameplay scenarios. In particular, additional Dark Pattern types and a more balanced comparison between the Victim and Exploiter perspectives could

provide further insight into how different forms of engagement influence awareness and reflection.

## 8 CONCLUSION

Overall, this study demonstrates the potential of role-based games as an educational approach for raising awareness of Dark Patterns. The findings suggest that the intervention supported participants' awareness of manipulative design practices and encouraged continued reflection several weeks after gameplay. However, the extent of change in immediate recognition and sustained vigilance cannot fully be established conclusively from the present measures. By embedding Dark Patterns within an interactive narrative, the game transformed abstract ethical concepts into more tangible and memorable experiences. The high rates of real-world recognition and social discussion reported weeks after the study further suggest that participants continued to reflect on manipulative interfaces beyond the gameplay experience itself.

The study highlights the value of perspective shifts in awareness-focused game design. While many existing educational approaches primarily focus on the user as a passive victim to teach them what Dark Patterns are, this often leaves users aware of the manipulation but frustrated and unsure of how to oppose it [8]. Complicity addresses this gap by exposing why and where these designs exist. By placing players in the Exploiter role, the game exposes them to the professional constraints and profit-driven goals that often take priority over user interests. [1] Players experience how Dark Patterns are deliberately created, justified and normalized within a simulated professional setting, which also illustrates the dangers of "role morality," demonstrating how professional contexts can encourage emotional distance from the ethical consequences of these decisions.[13]

By making the systemic pressures of designer complicity visible, the game may support a more structural understanding of Dark Patterns among both future practitioners and everyday users, while encouraging more critical engagement with digital systems. Overall, the findings suggest that role-based interactive experiences can serve as an effective approach for fostering awareness of manipulative design practices, particularly when they move beyond simple exposure and instead encourage reflection on the systems and motivations behind them.

## References

- [1] Gray et al. “The Dark Patterns Side of UX Design”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018. DOI: 10.1145/3173574.3174108.
- [2] Harry Brignull. *Deceptive Design*. 2010. URL: <https://www.deceptive.design/> (visited on 10/30/2025).
- [3] Mathur et al. “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*. 2021. DOI: 10.1145/3411764.3445610.
- [4] Gray et al. “An Ontology of Dark Patterns Knowledge: Foundations, Definitions, and a Pathway for Shared Knowledge-Building”. In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024. DOI: 10.1145/3613904.3642436.
- [5] Shoshana Zuboff. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs, 2019, pp. 14–16.
- [6] Meinhardt et al. “Mind Games! Exploring the Impact of Dark Patterns in Mixed Reality Scenarios”. In: *Proceedings of the ACM on Human-Computer Interaction* (2025). DOI: 10.1145/3743709.
- [7] Di Geronimo et al. “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*. 2020. DOI: 10.1145/3313831.3376600.
- [8] Bongard-Blanchy et al. ““I am Definitely Manipulated, Even When I am Aware of it. It’s Ridiculous!” – Dark Patterns from the End-User Perspective”. In: *Proceedings of the 2021 ACM Designing Interactive Systems Conference*. 2021. DOI: 10.1145/3461778.3462086.
- [9] Jamie Luguri and Lior Strahilevitz. “Shining a Light on Dark Patterns”. In: *Journal of Legal Analysis* 13 (2021), pp. 43–109. DOI: 10.1093/jla/laaa006.
- [10] Mathur et al. “Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites”. In: *Proceedings of the ACM on Human-Computer Interaction* (2019). DOI: 10.1145/3359183.
- [11] Kronhardt et al. “Trickery: Exploring a Serious Game Approach to Raise Awareness of Deceptive Patterns”. In: *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia*. 2024. DOI: 10.1145/3701571.3701588.
- [12] Lucas Pope. *Papers, Please*. Video game. Production: 3909. 2013.
- [13] Gibson. “Contrasting Role Morality and Professional Morality: Implications for Practice”. In: *Journal of Applied Philosophy* (2003). DOI: 10.1111/1468-5930.00232.
- [14] Grudpan et al. “Exploring the Effect of Game Premise in Cooperative Digital Board Games”. In: *1st Joint International Conference on Entertainment Computing and Serious Games*. 2019. DOI: 10.1007/978-3-030-34644-7\_17.
- [15] Chivukula et al. “Anti-Heroes: A Role-Based Method to Encourage Ethical Deliberation”. In: *Proceedings of the 2025 ACM Designing Interactive Systems Conference*. 2025. DOI: 10.1145/3715336.3735776.
- [16] All et al. “Assessing the Effectiveness of Digital Game-Based Learning: Best Practices”. In: *Computers & Education* (2016). DOI: 10.1016/j.compedu.2015.10.007.
- [17] Naul et al. “Why Story Matters: A Review of Narrative Storytelling in Serious Games”. In: *Journal of Educational Computing Research* 58.3 (2020), pp. 687–707. DOI: 10.1177/0735633119859904.
- [18] Bahrini et al. “Good vs. Evil: Investigating the Effect of Game Premise in a Smart Home Security Educational Game”. In: *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*. 2020. DOI: 10.1145/3383668.3419887.
- [19] Domínguez et al. “The Mimesis Effect: The Effect of Roles on Player Choice in Interactive Narrative Role-Playing Games”. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016, pp. 3438–3449. DOI: 10.1145/2858036.2858141.
- [20] Grizzard et al. “Being Bad in a Video Game Can Make Us More Morally Sensitive”. In: *Cyberpsychology, Behavior, and Social Networking* (2014). DOI: 10.1089/cyber.2013.0658.

- [21] Albert Bandura. *Manual for Constructing Moral Disengagement Scales*. URL: <https://albertbandura.com/albert-bandura-moral-disengagement.html>.
- [22] Mildner et al. “Defending Against the Dark Arts: Recognising Dark Patterns in Social Media”. In: *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 2023. DOI: 10.1145/3563657.3595964.
- [23] Niknejad et al. “Level Up or Game Over: Exploring How Dark Patterns Shape Mobile Games”. In: *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia*. 2024. DOI: 10.1145/3701571.3701604.

## Appendix

	Gameplay_Time	
	E	V
Valid	18	18
Missing	0	0
Mean (arithmetic)	21.39	18.22
Std. Deviation	7.005	6.620

### *Descriptive Statistics*

Session_Time	
Valid	36
Missing	0
Mean (arithmetic)	58.75
Std. Deviation	11.11

	Interview_Time	
	E	V
Valid	18	18
Mean (arithmetic)	16.17	13.67
Std. Deviation	5.009	4.446

Figure 15: Average gameplay-, interview-, and overall session duration.

**Additional study materials, including survey forms, raw data, and the game files related to Complicity, can be accessed at:**

<https://nc.uni-bremen.de/index.php/s/6DkPEg7G4WiFAS7>

### **Declaration on the Use of Generative AI**

Generative AI tools (GPT-4o, Gemini Image generation and Elevenlabs) were used to assist the development of this thesis and the accompanying game project *Complicity* for language revision, grammar and spelling checks, and the refinement of written passages. AI assistance was also used to help formulate theme names in the reflexive thematic analysis of the interviews, as well as for the creation of some visual game assets and the AI assistant voice audio which is referenced in the license document for third-party assets used within the games.

All AI-supported outputs were reviewed and edited independently by the author. The author takes responsibility for all content presented in this thesis.

# Interview Analysis - Codes

---

## Victim Condition

### Prior Awareness and Knowledge

Code	Description	Example	Frequency
Prior Concept Awareness	Participants knew about Dark Patterns before participating.	"I've seen videos in the past that explained it." (P. 21)	14
Limited or no Prior Awareness	Participants had very little or no knowledge of Dark Patterns beforehand.	"I never heard of it. I know that a lot of online shopping is misleading on purpose, but I didn't know that it was so organized, that there was so much behind it." (P. 29)	4

### Emotional / Psychological Impact

Code	Description	Example	Frequency
Frustration	Participants were frustrated or annoyed by the Dark Patterns in the game.	"I remember things best, when I rage about them and this thing really made me rage." (P. 17)	6
Resignation / Sense of Helplessness	Their awareness is coupled with a feeling that these patterns are unavoidable/standard.	"A lot of patterns like the subscription one is really intentionally bad but you can't do anything about it. It should be better regulated." (P. 21)	5

### Awareness & Learning

<b>Code</b>	<b>Description</b>	<b>Example</b>	<b>Frequency</b>
Recognition of Manipulation	Participants noticed the manipulative intent of in-game interfaces or framing.	“The Roach Motel left the biggest impression. It’s the worst Dark Pattern to encounter in real-life.” (P. 13)	18
Awareness of "False Choice"	Participants noted that even when the game presented an option to decline it wasn't a "real" choice.	“I cancelled the update and it kept beeping which was really funny that it just stops to work which in my experience happens irl too. They give you the option to not update but then it doesn’t work anymore. So it’s not really an option.” (P. 21)	7
Real-World Relatability	Participants drew the connection between game tasks and personal negative experiences.	“Canceling your Amazon Prime subscription. They have five steps before you can actually cancel it and they say you will miss out on this and that.” (P. 7)	17
Heightened Vigilance	Participants state that they will be more careful with interfaces after playing.	“I was already very aware of it, but the game gave me another perspective. Now I know I need to focus on it even more to avoid getting scammed.” (P. 3)	14

### Role Preferences

<b>Code</b>	<b>Description</b>	<b>Example</b>	<b>Frequency</b>
Prefers Designer Role	Participants believe the designer role would be better for learning about Dark Patterns.	“I think the designer version would be better for teaching because you get more understanding of how they created it.” (P. 9)	7
Prefers Victim Role	Participants feel that playing the victim role is more memorable.	“I remember things best when I rage about them, and this thing really made me rage. ” (P. 17)	6

Undecided	They think both roles are valuable to raise awareness about the topic.	"Either way I think you'd learn from the games." (P. 13)	5
-----------	--	--	---

## Exploiter Condition

### Prior Awareness and Knowledge

Code	Description	Example	Frequency
Prior Concept Awareness	Participants knew about Dark Patterns before participating.	"I already did research on this topic myself for school." (P. 10)	16
Limited or no Prior Awareness	Participants had very little or no knowledge of Dark Patterns beforehand.	"I didn't know the term. I guess it's about deceptive marketing strategies." (P. 20)	2

### Emotional / Psychological Impact

Code	Description	Example	Frequency
Complicit Guilt	Feelings of guilt or "feeling evil" by being the agent of harm.	"Actively trying to scam the costumer was the worst part for me in the exploiter gameplay. Doing this to other people makes you self reflect and think about what your actions cause." (P. 18)	5
Pragmatic Immersion	Compliance Immersed in the professional role; making choices based on game instructions or perceived "best" design logic without strong	"I think I just followed through with the instructions that were given to me and in that sense I think I did good." (P. 32)	10

	emotional attachment.		
Amusement	Finding the "evil" roleplay fun or entertaining rather than stressful.	"I was a little bit surprised by how fun I found it to think from the other side. Like, how can I manipulate people into having the worst experience possible while staying in the legal boundaries? I actually found that interesting and it entertained me." (P. 4)	3

### Awareness and Learning

<b>Code</b>	<b>Description</b>	<b>Example</b>	<b>Frequency</b>
Recognition of Manipulative Design	Participants described the manipulative intent of in-game interfaces.	"I really disliked the interface options for Mochi Garden that were manipulative." (P. 2)	8
Recognition of Manipulative Framing	Participants noticed the pressuring and manipulative task framing.	"The tasks are framed very cute while you actually manipulate people." (P. 4)	9
Revealed Mechanisms	Participants felt like "peeking behind the curtains" of where, why and how Dark Patterns are implemented.	"I thought it was really cool to firsthand implement these in order to get more of an idea of how they work and where they are integrated." (P. 12)	11
Real-World Relatability	Participants drew the connection between game tasks and personal negative experiences.	"I think definitely cookies, because it's very difficult to just decline them. The options are hidden, and some things are sometimes pre-marked." (P. 12)	18
Heightened	Participants state	"Definitely, because I've never	16

Vigilance	that they will be more careful with interfaces after playing.	thought about being the person that is designing these things. Because when you learn about dark patterns, you're like, "Oh, how could anybody ever use these when actually working in marketing?" (P. 36)	
-----------	---	--	--

### Reflection on the other Role

Code	Description	Example	Frequency
Prefers Designer Role	Participants believe the designer role would be better for learning about Dark Patterns.	"Understanding comes from seeing the other side (designer) and forcing the people to embody the thing they are not aware, making them do the things that normally happen to them." (P. 14)	11
Anticipated Role-Based Frustration	Participants express hypothetical annoyance or frustration when imagining themselves playing the victim gameplay.	"I'd be extremely annoyed in the victim situation because every website is evil and plays tricks on you [...] while the victim role is valuable, I can experience that just by going on the next website." (P. 8)	10
Prefers Victim Role	Participants feel that playing the victim role is more memorable.	"I do think it's more effective 'cause it causes stronger emotional reactions." (P. 30)	4
Undecided	They think both roles are valuable to raise awareness about the topic.	"I think to have the full picture of it, both sides are important to know how it actually works and what options there are. I don't have a favorite." (P. 22)	3

## Hinweise zu den offiziellen Erklärungen

1. Die folgende Seite mit den offiziellen Erklärungen

A) Eigenständigkeitserklärung

B) Erklärung zur Veröffentlichung von Bachelor- oder Masterarbeiten

C) Einverständniserklärung über die Bereitstellung und Nutzung der Bachelorarbeit / Masterarbeit in elektronischer Form zur Überprüfung durch eine Plagiatssoftware

ist entweder direkt in jedes Exemplar der Bachelor- oder Masterarbeit fest mit einzubinden oder unverändert im Wortlaut in jedes Exemplar der Bachelor- oder Masterarbeit zu übernehmen.

**Bitte achten Sie darauf, jede Erklärung in allen drei Exemplaren der Arbeit zu unterschreiben.**

2. In der digitalen Fassung kann auf die Unterschrift verzichtet werden. Die Angaben und Entscheidungen müssen jedoch enthalten sein.

### Zu B)

Die Einwilligung kann jederzeit durch Erklärung gegenüber der Universität Bremen, mit Wirkung für die Zukunft, widerrufen werden.

### Zu C)

Das Einverständnis der dauerhaften Speicherung des Textes ist freiwillig.

Die Einwilligung kann jederzeit durch Erklärung gegenüber der Universität Bremen, mit Wirkung für die Zukunft, widerrufen werden.

Weitere Informationen zur Überprüfung von schriftlichen Arbeiten durch die Plagiatsoftware sind im Nutzungs- und Datenschutzkonzept enthalten. Diese finden Sie auf der Internetseite der Universität Bremen.

---

## Notes on the official declarations

1. The following pages with the official declarations

A) Declaration of Authorship

B) Declaration on the Publication of Bachelor's or Master's Thesis

C) Declaration of Consent for the Provision and Use of the Bachelor's Thesis / Master's Thesis in Electronic Form for Review by Plagiarism Software

is to be either integrated directly into each copy of the bachelor's or master's thesis or adopted unchanged in the wording of each copy of the bachelor's or master's thesis.

**Please be sure to sign each declaration in all three copies of the thesis.**

2. The signature can be omitted from the digital version. However, the information and decisions must be included.

### Regarding B)

The consent can be revoked at any time with future effect by notifying the University of Bremen.

### Regarding C)

Consent for the permanent storage of the text is voluntary. The consent can be revoked at any time with future effect by notifying the University of Bremen.

Further information on the checking of written work using plagiarism software can be found in the data protection and usage concept. This can be found on the University of Bremen website.

## **A) Eigenständigkeitserklärung / Declaration of Authorship**

Ich versichere, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Teile meiner Arbeit, die wortwörtlich oder dem Sinn nach anderen Werken entnommen sind, wurden unter Angabe der Quelle kenntlich gemacht. Gleiches gilt auch für Zeichnungen, Skizzen, bildliche Darstellungen sowie für Quellen aus dem Internet, dazu zählen auch KI-basierte Anwendungen oder Werkzeuge.  
Die Arbeit wurde in gleicher oder ähnlicher Form noch nicht als Prüfungsleistung eingereicht.

I hereby affirm that I have written the present work independently and have used no sources or aids other than those indicated. All parts of my work that have been taken from other works, either verbatim or in terms of meaning, have been marked as such, indicating the source. The same applies to drawings, sketches, pictorial representations and sources from the Internet, including AI-based applications or tools. The work has not yet been submitted in the same or a similar form as a final examination paper.

Ich habe KI-basierte Anwendungen und/oder Werkzeuge genutzt und diese im Anhang "Nutzung KI basierte Anwendungen" dokumentiert.

I have used AI-based applications and/or tools and documented them in the appendix "Use of AI-based applications".

**B) Erklärung zur Veröffentlichung von Bachelor- oder Masterarbeiten**  
Declaration regarding the publication of bachelor's or master's thesis

Die Abschlussarbeit wird zwei Jahre nach Studienabschluss dem Archiv der Universität Bremen zur dauerhaften Archivierung angeboten. Archiviert werden:

Two years after graduation, the thesis is offered to the archive of the University of Bremen for permanent archiving. The following are archived:

- 1) Masterarbeiten mit lokalem oder regionalem Bezug sowie pro Studienfach und Studienjahr 10 % aller Masterarbeiten  
Master's theses with a local or regional focus, as well as per subject and academic year 10% of all Master's thesis
- 2) Bachelorarbeiten des jeweils ersten und letzten Bachelorabschlusses pro Studienfach und Jahr.  
Bachelor's thesis for the first and last bachelor's degrees per subject and year.

Ich bin damit einverstanden, dass meine Abschlussarbeit im Universitätsarchiv für wissenschaftliche Zwecke von Dritten eingesehen werden darf.  
I agree that my thesis may be viewed by third parties in the university archive for academic purposes.

Ich bin damit einverstanden, dass meine Abschlussarbeit nach 30 Jahren (gem. §7 Abs. 2 BremArchivG) im Universitätsarchiv für wissenschaftliche Zwecke von Dritten eingesehen werden darf.  
I agree that my thesis may be viewed by third parties for academic purposes in the university archive after 30 years (in accordance with §7 para. 2 BremArchivG).

Ich bin **nicht** damit einverstanden, dass meine Abschlussarbeit im Universitätsarchiv für wissenschaftliche Zwecke von Dritten eingesehen werden darf.  
I do not consent to my thesis being made available in the university archive for third parties to view for academic purposes.

### **C) Einverständniserklärung zur elektronischen Überprüfung der Arbeit auf Plagiate Declaration of consent for electronic checking of the work for plagiarism**

Eingereichte Arbeiten können nach § 18 des Allgemeinen Teil der Bachelor- bzw. der Masterprüfungsordnungen der Universität Bremen mit qualifizierter Software auf Plagiatsvorwürfe untersucht werden.

Zum Zweck der Überprüfung auf Plagiate erfolgt das Hochladen auf den Server der von der Universität Bremen aktuell genutzten Plagiatssoftware.

Submitted papers can be checked for plagiarism using qualified software in accordance with § 18 of the General Section of the Bachelor's or Master's Degree Examination Regulations of the University of Bremen. For the purpose of checking for plagiarism, the upload to the server is done using the plagiarism software currently used by the University of Bremen.

Ich bin damit einverstanden, dass die von mir vorgelegte und verfasste Arbeit zum oben genannten Zweck dauerhaft auf dem externen Server der aktuell von der Universität Bremen genutzten Plagiatssoftware, in einer institutionseigenen Bibliothek (Zugriff nur durch die Universität Bremen), gespeichert wird.

I agree that the work I have submitted and written will be stored permanently on the external server of the plagiarism software currently used by the University of Bremen, in a library belonging to the institution (accessed only by the University of Bremen), for the above-mentioned purpose.

Ich bin **nicht** damit einverstanden, dass die von mir vorgelegte und verfasste Arbeit zum o.g. Zweck dauerhaft auf dem externen Server der aktuell von der Universität Bremen genutzten Plagiatssoftware, in einer institutionseigenen Bibliothek (Zugriff nur durch die Universität Bremen), gespeichert wird.

I do not consent to the work I submitted and wrote being permanently stored on the external server of the plagiarism software currently used by the University of Bremen, in a library belonging to the institution (accessed only by the University of Bremen), for the above-mentioned purpose.

Das Einverständnis der dauerhaften Speicherung des Textes ist freiwillig. Die Einwilligung kann jederzeit durch Erklärung gegenüber der Universität Bremen, mit Wirkung für die Zukunft, widerrufen werden. Weitere Informationen zur Überprüfung von schriftlichen Arbeiten durch die Plagiatsoftware sind im Nutzungs- und Datenschutzkonzept enthalten. Diese finden Sie auf der Internetseite der Universität Bremen.

Consent to the permanent storage of the text is voluntary. Consent can be withdrawn at any time by making a declaration to this effect to the University of Bremen, with effect for the future. Further information on the checking of written work using plagiarism software can be found in the data protection and usage concept. This can be found on the University of Bremen website.

Mit meiner Unterschrift versichere ich, dass ich die obenstehenden Erklärungen gelesen und verstanden habe und bestätige die Richtigkeit der gemachten Angaben.

With my signature, I confirm that I have read and understood the above explanations and confirm the accuracy of the information provided.

Datum / Date

Unterschrift/ Signature