

The Impact of Dark Patterns in E-Commerce on User Behaviour and Decision-Making

Bijelac, Domagoj

Master's thesis / Diplomski rad (sveučilišni)

2025

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Rijeka, - FACULTY OF ENGINEERING / Sveučilište u Rijeci, - Tehnički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:190:577950>

Rights / Prava: [Imenovanje 4.0 međunarodna](#) / [Attribution 4.0 International](#)

Download date / Datum preuzimanja: **2026-05-02**



Repository / Repozitorij:

[Repository of the University of Rijeka, Faculty of Engineering / Repozitorij Tehničkog fakulteta u Rijeci](#)



Dabar

UNIVERSITY OF RIJEKA
FACULTY OF ENGINEERING
Graduate University Study of Computing

Master's thesis

**The Impact of Dark Patterns in
E-Commerce on User Behaviour and
Decision-Making**

Rijeka, September 2025.

Domagoj Bijelac
0069088384

UNIVERSITY OF RIJEKA
FACULTY OF ENGINEERING
Graduate University Study of Computing

Master's thesis

**The Impact of Dark Patterns in
E-Commerce on User Behaviour and
Decision-Making**

Mentor: prof.dr.sc. Sandi Ljubić

Rijeka, September 2025.

Domagoj Bijelac
0069088384

Rijeka, 11.03.2025.

Zavod: Zavod za računarstvo
Predmet: Interakcija čovjeka i računala

ZADATAK ZA DIPLOMSKI RAD

Pristupnik: **Domagoj Bijelac (0069088384)**
Studij: Sveučilišni diplomski studij računarstva (1400)
Modul: Programsko inženjerstvo (1441)

Zadatak: **Utjecaj tamnih obrazaca u e-trgovini na ponašanje korisnika i donošenje odluka / The Impact of Dark Patterns in E-Commerce on User Behaviour and Decision-Making**

Opis zadatka:

Cilj rada je, kroz empirijsko istraživanje, analizirati utjecaj tamnih obrazaca u e-trgovini na korisničko ponašanje i proces donošenja odluka. Potrebno je implementirati dvije inačice e-trgovine: aplikaciju s uključenim tamnim obrascima te aplikaciju zasnovanu na etičkom dizajnu sučelja. Kod prve je iskustvo kupovine optimizirano za angažman korisnika, prikupljanje podataka i maksimizaciju profita, korištenjem manipulativnih dizajnerskih taktika poput skrivenih troškova, prisilnih pretplata i vremenskog pritiska. Nasuprot tome je transparentno i korisniku usmjereno iskustvo bez manipulativnih taktika, s jasnim prikazom svih relevantnih informacija. Potrebno je organizirati i provesti HCI eksperiment sa stvarnim korisnicima te definirati i analizirati ključne metrike interakcije s aplikacijama, uključujući kako objektivne pokazatelje tako i subjektivni dojam korisnika.

Rad mora biti napisan prema Uputama za pisanja diplomskih / završnih radova koje su objavljene na mrežnim stranicama studija.

Zadatak uručen pristupniku: 21.03.2025.

Mentor:
prof. dr. sc. Sandi Ljubić

Predsjednik povjerenstva za
diplomski ispit:
prof. dr. sc. Miroslav Joler

Statement on independent creation of thesis

I declare that I created this thesis independently.

Rijeka, September 2025.

Domagoj Bijelac

Acknowledgement

I would like to express my sincere gratitude to my mentor prof. dr. sc. Sandi Ljubić for his invaluable mentorship, insights, and guidance. I am deeply thankful to my family and my partner for their unwavering and everlasting support, and to the friends I made during my studies, with whom I shared knowledge, overcame challenges, and built lasting friendships.

Contents

List of Figures	ix
List of Tables	xi
1 Introduction	1
2 Conceptual foundations and related work	4
2.1 A taxonomy of implemented dark patterns	4
2.1.1 Urgency and scarcity	4
2.1.2 Misdirection	5
2.1.3 Social proof	5
2.1.4 Sneaking	5
2.1.5 Obstruction	6
2.2 The ethical counterpoint: User-centered design	6
2.3 Measuring the impact on user behavior	7
3 Methodology	8
3.1 Research design	8
3.1.1 Independent and dependent variables	8
3.2 Participants	11
3.2.1 Recruitment and study setup	11

Contents

3.2.2	Random assignment and final sample	12
3.3	The experimental application	14
3.3.1	Technology stack and architecture	14
3.3.2	Version A: The dark pattern interface	16
3.3.3	Version B: The ethical design interface	19
3.3.4	Side-by-side interface comparison	20
3.4	Experimental procedure	27
3.4.1	Pre-task experience and onboarding	27
3.4.2	Shopping experience and data collection	28
3.4.3	Post-experiment survey and completion	31
3.5	Data analysis	32
3.5.1	Quantitative analysis	32
3.5.2	Qualitative (survey) analysis	33
3.5.3	Advanced and demographic analyses	33
4	Results	34
4.1	Quantitative findings	34
4.1.1	Spending metrics	34
4.1.2	Time & engagement metrics	38
4.1.3	Discrete & binary metrics	41
4.2	Qualitative findings	46
4.2.1	Perceived ease & clarity of use	46
4.2.2	Perceived control & trust	47
4.2.3	Awareness of manipulative elements	48
4.3	Advanced & demographic analyses	50
4.3.1	Impact of awareness on behavior (Group A)	50
4.3.2	Demographic correlation and segmentation	51

Contents

4.3.3	Shame modal effectiveness (Group A)	59
5	Discussion	60
5.1	Interpretation of key findings	60
5.2	Implications and recommendations	63
5.3	Study limitations	63
5.4	Future work	64
6	Conclusion	65
	Bibliography	67
	Abstract	69

List of Figures

3.1	Study configuration details as presented on the Prolific platform. . .	12
3.2	Demographic distribution of the final participant sample (N=60). . .	13
3.3	A Misdirection modal that combines a Confirmshaming message with Visual interference, where the business-preferred option ("Yes, Keep My Protection") is styled as the primary action.	17
3.4	A product card in Version A featuring multiple scarcity and urgency patterns: an "Exclusive Offer" tag, a low stock message, and a count-down timer.	18
3.5	A social proof pop-up visible in the lower left corner of the product detail modal, designed to imply popularity and encourage a purchase.	19
3.6	Comparison of the home page.	21
3.7	Comparison of the shopping cart.	22
3.8	Comparison of the checkout page.	24
3.9	The two-step post-checkout journey in Version A, showing the confirmshame modal.	25
3.10	The single-step post-checkout journey in Version B.	26
3.11	The four key screens of the standardized onboarding procedure. . . .	28
4.1	Comparison of Final order value (cleaned data).	35
4.2	Comparison of Subtotal (cleaned data).	36
4.3	Comparison of Cost discrepancy (cleaned data).	37

List of Figures

4.4	Comparison of Total session time (cleaned data).	38
4.5	Comparison of Time to first engagement (cleaned data).	39
4.6	Comparison of Passive browsing time (cleaned data).	40
4.7	Passively viewed products.	42
4.8	Products clicked.	42
4.9	Registration rate.	43
4.10	Checkout protection rate.	43
4.11	Warranty purchase rate.	44
4.12	Saved card rate.	44
4.13	Provided birthday rate.	45
4.14	Provided interests rate.	45
4.15	Perceived ease & clarity of use scores (Q1 & Q2).	47
4.16	Perceived control & trust scores (Q3 & Q4).	48
4.17	Awareness of manipulative elements scores (Q5 & Q6).	49
4.18	Correlation between Awareness score and Final order value (full dataset).	51
4.19	Correlation between Age and Final order value (cleaned data).	52
4.20	Correlation between Age and Perceived control score (q3).	53
4.21	Correlation between Age and Website trust score (q4).	54
4.22	Correlation between Age and Feeling of pressure score (q5).	55
4.23	Correlation between Age and Awareness of manipulation score (q6).	56
4.24	Comparison of Mean scores for Perceived control and Trust by gender.	58
4.25	Comparison of Mean scores for Feeling of pressure and Awareness by gender.	58
4.26	Effectiveness of Shame modals in Group A.	59

List of Tables

4.1	Descriptive statistics for the Final Order Value, comparing Group A (Dark) and Group B (Ethical) across both the full and cleaned datasets.	35
4.2	Descriptive statistics for the Subtotal, comparing Group A and Group B across both the full and cleaned datasets.	36
4.3	Descriptive statistics for the Cost Discrepancy, comparing Group A and Group B across both the full and cleaned datasets.	37
4.4	Descriptive statistics for Total session time, comparing Group A and Group B across both the full and cleaned datasets.	38
4.5	Descriptive statistics for Time to first engagement, comparing Group A and Group B across both the full and cleaned datasets.	39
4.6	Descriptive statistics for Passive browsing time, comparing Group A and Group B across both the full and cleaned datasets.	40
4.7	Descriptive statistics for Passively viewed products.	41
4.8	Descriptive statistics for Products clicked.	41
4.9	Descriptive statistics for Registration rate.	42
4.10	Descriptive statistics for Checkout protection rate.	43
4.11	Descriptive statistics for Warranty purchase rate.	43
4.12	Descriptive statistics for Saved card rate.	44
4.13	Descriptive statistics for Provided birthday rate.	45
4.14	Descriptive statistics for Provided interests rate.	45

List of Tables

4.15	Descriptive statistics for Q1: Website easy to navigate and understand.	46
4.16	Descriptive statistics for Q2: Clarity about total cost.	46
4.17	Descriptive statistics for Q3: Feeling in control of decisions.	47
4.18	Descriptive statistics for Q4: Trust in website transparency.	48
4.19	Descriptive statistics for Q5: Feeling pressured to purchase.	48
4.20	Descriptive statistics for Q6: Awareness of manipulative design elements.	49
4.21	Descriptive statistics for Awareness analysis (Group A).	50
4.22	Descriptive statistics for Age and Final order value.	52
4.23	Descriptive statistics for Age and Perceived control (q3).	53
4.24	Descriptive statistics for Age and Website trust (q4).	54
4.25	Descriptive statistics for Age and Feeling of pressure (q5).	55
4.26	Descriptive statistics for Age and Awareness of manipulation (q6). .	56
4.27	Descriptive statistics for Perceived control (q3) and Trust (q4) by gender.	57
4.28	Descriptive statistics for Pressure (q5) and Awareness (q6) by gender.	57
4.29	Descriptive statistics for Shame modal user outcomes (Group A). . .	59

Chapter 1

Introduction

As e-commerce continues to grow, user interface and user experience (UI/UX) design has become essential to achieve success in the digital marketplace. The role of UI/UX is critical for maintaining user engagement and ensuring long-term sustainability, enabling platforms to remain competitive by enhancing user satisfaction, attracting more users, and boosting profitability [1]. A well-designed UI makes a website easier to use and more interactive, which in e-commerce, is a key driver of customer satisfaction. Beyond service and product quality, users value a high-quality user experience, which encompasses functionality, usability, and content [2].

Within this context, persuasive design has become a central practice. Persuasive design is a method that uses the design of products and services to encourage users to behave in ways that achieve a particular goal [3]. Since users are often concerned about security, privacy and trust when shopping online, persuasive principles can be implemented ethically to guide users through complex checkout processes or help them find relevant products. When designed and implemented correctly, persuasive design can turn browsing users into purchasing customers, contributing positively to e-commerce success [4]. This study was motivated by the need to further research persuasive website design from the user's perspective, particularly the fine line where persuasion turns into manipulation.

In e-commerce, the line between persuasion and deception is often crossed with the use of **dark patterns**. First defined by Harry Brignull, dark patterns are inter-

face elements that intentionally guide, pressure or mislead users into decisions that are against their own interests but profitable for the service provider [5]. If users were fully informed and aware of alternatives, they might choose differently. These interface designs are a prevalent practice on digital platforms, found on thousands of shopping websites, social media platforms and mobile apps [6].

Dark patterns can cause frustration and inconvenience, but they can also be actively deceiving and harmful, potentially leading to financial loss, involuntary disclosure of personal data or the development of compulsive behaviors [6]. While their prevalence is well-documented, a critical gap remains in the quantitative understanding of their specific, measurable impact on user behavior in a controlled experimental setting. It is often assumed these patterns increase revenue and data collection, but their cost in terms of user friction, session time and subjective trust is less understood. This study aims to address this gap by directly comparing an e-commerce interface built with dark patterns against an ethical counterpart to measure the precise effects on both objective user actions and subjective user perceptions.

To investigate the problem, this thesis seeks to answer the following research questions:

- How does an e-commerce interface implementing dark patterns affect users objective, measurable behaviors, specifically metrics such as: final order value, total session time and conversion rates on optional items (e.g., warranties, account registration) compared to an ethical interface?
- What is the impact of dark patterns on users subjective perceptions, particularly their sense of control over their decisions, their trust in the website, and their perceived ease of use?
- To what extent are users consciously aware of the manipulative design elements they encounter, and does this awareness correlate with a change in their purchasing behavior?

The scope of this research is confined to a controlled human computer interface (HCI) experiment conducted within a purpose-built, simulated e-commerce application named "Shop-Mania." The study focuses on a set of dark patterns drawn from the categories of urgency, scarcity, misdirection, social proof, sneaking, and obstruc-

Chapter 1. Introduction

tion. The findings are based on data collected from 60 participants recruited via the Prolific academic research platform, who were required to complete the task on a desktop device. The study doesn't analyze the long-term effects of dark patterns on brand loyalty, nor does it explore their impact on mobile interfaces or other cultural contexts outside of the recruited demographic.

The remainder of this thesis is structured as follows: **Chapter 2** provides a review of the relevant literature, establishing a conceptual foundation by defining the specific dark patterns implemented and contrasting them with principles of ethical design. **Chapter 3** details the experimental methodology, including the research design, participant recruitment, the technical architecture of the application, and the step-by-step experimental procedure. **Chapter 4** presents the quantitative and qualitative results of the experiment, using statistical tests to compare the two user groups. **Chapter 5** discusses the interpretation and implications of these results, addresses the study's limitations, and suggests directions for future work. Finally, **Chapter 6** provides a summary of the findings and concludes the thesis.

Chapter 2

Conceptual foundations and related work

2.1 A taxonomy of implemented dark patterns

Now that the problem of dark patterns has been established in a broader context, this chapter provides a detailed review of the specific patterns implemented in the experimental application. To ensure that the study is grounded in established research, the classification of these patterns primarily follows the empirically derived taxonomy from Mathur et al.’s large-scale analysis of 11,000 shopping websites [6]. This framework establishes a basis for understanding the mechanisms inside the manipulative interface.

2.1.1 Urgency and scarcity

These patterns leverage the **scarcity bias**, a cognitive heuristic where individuals place a higher value on items they believe are rare or available for a limited time [7]. **Urgency** creates time pressure by imposing a deadline. In this study, this was implemented through *countdown timers* on product cards, the product detail modal, and the checkout page, a tactic Mathur et al. found on hundreds of major websites [6].

Scarcity suggests limited product availability. This was implemented via *low-stock messages* (e.g., "Only 5 left in stock!"), the single most frequent pattern observed by Mathur et al. [6].

2.1.2 Misdirection

Misdirection patterns use deceptive visual cues and emotional language to steer users. **Confirmshaming** employs guilt to manipulate a user's choice, leveraging the **framing effect** [8]. In this experiment, it was used extensively in modals that appeared when a user tried to opt-out of a pre-selected choice, such as removing a warranty or saving their card details. The language was designed to make the user feel their decision was risky or unwise.

Visual interference was paired with confirmshaming. The business-preferred option (e.g., "Keep My Protection") was styled as a prominent, often green, primary button, while the user's intended action was a de-emphasized, neutral button. This pattern increases cognitive load and exploits common user tendencies to click the most prominent action [6].

2.1.3 Social proof

This category leverages the principle of conformity, suggesting that an action is more correct if others are performing it [7]. **Activity messages** were implemented in via a pop-up on the product detail modal, which claimed that another person had recently purchased the viewed item. This pattern, creates a false sense of demand and popularity.

2.1.4 Sneaking

Sneaking involves adding items or charges to a user's order without their explicit consent. **Sneak into basket** was implemented by pre-selecting an optional product warranty for every item added to the cart. This shifts the burden to the user to notice and opt-out.

Hidden costs (Drip pricing) was implemented with a mandatory "processing fee" that was only revealed on the final checkout page. This practice exploits the **sunk cost fallacy**; by the time the fee is revealed, the user has already invested significant effort and is less likely to abandon the purchase [6].

2.1.5 Obstruction

This pattern makes a process more difficult than it needs to be. This was employed by **randomizing the product grid** on every page load, obstructing the user's ability to easily relocate a previously seen item, forcing more scrolling and scanning.

2.2 The ethical counterpoint: User-centered design

In direct opposition to dark patterns, the field of HCI has long established principles for a user-centered design, which prioritize the user's needs, goals, and autonomy.

According to industry leaders like the Nielsen Norman Group, a usable and ethical interface must be learnable, efficient, and subjectively pleasant [9]. Central to these goals are the following principles, embedded in the ethical version:

- **Transparency:** All costs, including processing and shipping fees, were presented clearly and upfront in the shopping cart. There were no financial surprises.
- **User control and freedom:** All optional choices, such as warranties and saving payment details, were presented as opt-in and unchecked by default. Users could navigate and reverse actions without encountering shame modals or friction.
- **Clarity and honesty:** The interface communicated information factually, without false urgency, misleading social proof, or emotionally manipulative language. The product grid was stable and predictable, respecting the user's time and effort.

2.3 Measuring the impact on user behavior

To quantify the effects of different design paradigms, HCI research relies on a combination of objective and subjective metrics, providing a holistic view by capturing users actions and how they felt performing them.

Objective quantitative metrics — These are direct, numerical measurements of user interaction. The metrics collected in this study, such as: Final order value, Total session time, and the conversion rates for binary choices like the Registration rate, are standard quantitative measures used to assess task success and efficiency.

Subjective qualitative metrics — These metrics capture user’s perceptions and feelings. They are crucial for understanding concepts like trust, satisfaction, and perceived control, which objective data alone cannot reveal. The most common method for collecting this data is through post-task surveys employing a **Likert scale** [10]. This study uses a 5-point Likert scale survey to assess participants perceptions of the site’s ease of use, transparency, and trustworthiness, providing the critical qualitative data needed to answer the second and third research questions.

Chapter 3

Methodology

3.1 Research design

The study employed an independent **between-subjects experimental design**, also commonly known as A/B testing. In this design, participants were randomly assigned to one of two distinct groups, with each group interacting with only one version of the experimental application. This approach was chosen to prevent learning and carryover effects, which would be a significant risk in a within-subjects design where participants interact with both versions. If a participant was exposed to the ethical interface first, their subsequent interaction with the dark pattern interface would be biased, as the manipulative elements would now particularly stand out. Conversely, prior exposure to the dark pattern version could instill a sense of distrust that would unfairly influence their perception of the ethical design. The between-subjects design ensures that each participant's experience is independent, providing a reliable comparison of the two interfaces.

3.1.1 Independent and dependent variables

The experimental design was structured around a single independent variable and a comprehensive set of dependent variables, capturing objective user behavior and subjective perceptions.

Independent variable

The sole independent variable was the **user interface type**, which consisted of two levels:

- **Version A:** An e-commerce interface featuring a curated set of dark patterns.
- **Version B:** An e-commerce interface based on transparent and ethical design principles.

Dependent variables

A range of quantitative and qualitative metrics were collected to measure the effects of the independent variable.

The **quantitative dependent variables** captured objective, measurable user actions and outcomes:

- **Final order value:** The total monetary value of the user's completed transaction, including all fees, warranties and discounts.
- **Subtotal:** The value of the items in the cart only, not including user-selected add-ons like warranties.
- **Cost discrepancy:** A derived metric calculated as $(Final\ order\ value - Subtotal)$ to quantify the impact of hidden fees or discounts.
- **Total session time:** The total duration from the participant's successful login to their arrival at the post-experiment survey page.
- **Time to first engagement:** The time elapsed from the homepage loading to the user's first click on a product to view its details.
- **Passive browsing time:** The total time spent on the home page where the primary activity is browsing.
- **Active engagement time:** The total time spent on pages requiring active user input, specifically the product detail modal.
- **Passively viewed products:** The count of unique products that were at least 50% visible in the user's viewport.

- **Products clicked:** The count of unique products a user clicked to view in detail.
- **Registration rate:** A measure of whether a user registered or not.
- **Checkout protection rate:** A boolean measure of whether the user purchased the optional order protection at checkout.
- **Warranty purchase rate:** The number of optional warranties purchased, relative to the number of items in the cart.
- **Saved card rate:** A boolean measure of whether the user opted to save their payment details.
- **Provided birthday rate:** A boolean measure of whether the user provided their birthday, during registration in the optional information section.
- **Provided interests rate:** A boolean measure of whether the user provided their personal interests, during registration in the optional information section.

The **qualitative dependent variables** were collected via a mandatory post-experiment survey, using a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree):

- **Perceived ease & clarity of use**
 - The website was easy to navigate and understand.
 - I always felt clear about the total cost of my order before the final payment step.
- **Perceived control & trust**
 - I felt in complete control of my decisions while shopping.
 - I trust this website to be transparent and act in my best interest.
- **Awareness of manipulative elements**
 - I felt pressured to make a purchase or add extra items.
 - I noticed design elements that seemed intended to influence my purchasing decisions.

3.2 Participants

Participants for this study were recruited using **Prolific**, an online platform designed for academic research. Prolific was chosen for several key reasons: it is known for providing quality and reliable data from a diverse participant pool; it offers prescreening filters to ensure compliance with study requirements; it includes a built-in mechanism for securely distributing unique login credentials to each participant, which was essential for this experiment. Lastly, it also includes an option to restrict access to the experiment based on the participant’s device, in this case the experiment could be accessed from a desktop-only device.

3.2.1 Recruitment and study setup

The study recruited a total of **60 participants**, with an equal distribution of 30 participants per experimental group. A sample size of $N=30$ per group is consistent with widely cited quantitative guidelines in HCI research. While the Nielsen Norman Group recommends approximately 40 participants for quantitative usability studies [11], 30 is generally considered a defensible lower bound for detecting meaningful experimental effects, particularly when practical constraints such as budget limit larger sample sizes.

The study was published on Prolific with the neutral title and description shown in Figure 3.1 to avoid priming participants or revealing the true nature of the research. Participants were required to complete the study on a desktop device to have a consistent user experience. The geographical scope was limited to countries within the European Union and the United Kingdom. No other demographic restrictions, such as age or gender, were applied. The estimated completion time was set to 7 minutes, based on a median completion time of 5 minutes and 37 seconds observed during a pilot study. The maximum time allowed was 36 minutes. Participants were compensated at a rate of £6.00/hour, receiving £0.70 for their participation, in line with Prolific’s ethical reward standards.

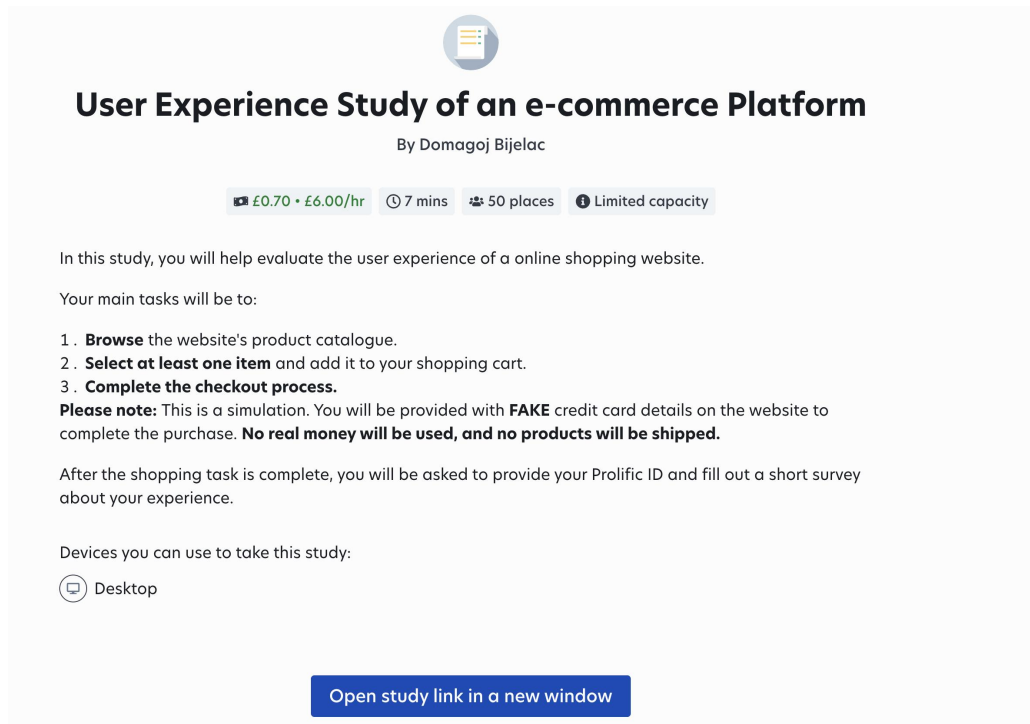


Figure 3.1 Study configuration details as presented on the Prolific platform.

3.2.2 Random assignment and final sample

Upon logging into the application, participants were randomly assigned to either Group A (Dark pattern) or Group B (Ethical design) by the server-side logic. The system was designed to maintain a balance between the groups by tracking the current count for each and assigning new users to the smaller group.

A total of 61 submissions were collected to reach the target of 60 approved participants. One participant's submission was rejected due to highly anomalous behavior. Prolific flagged the submission for manual review as it was completed in 2 minutes and 30 seconds, well below the expected time. An analysis of their collected metrics revealed that only a single product was passively viewed, suggesting the participant either used a mobile device against the study's rules or resized their browser window to be exceptionally small, preventing meaningful engagement with the task. Following this exclusion, another participant was recruited at random. The final sample consisted of 60 participants. Due to the random nature of the final assignment and

the single exclusion, the final group distribution was slightly imbalanced:

- **Group A (Dark pattern):** 29 participants
- **Group B (Ethical design):** 31 participants

A breakdown of key demographic characteristics used in Demographic analysis (Section 4.3) of the final participant sample is provided in Figure 3.2.

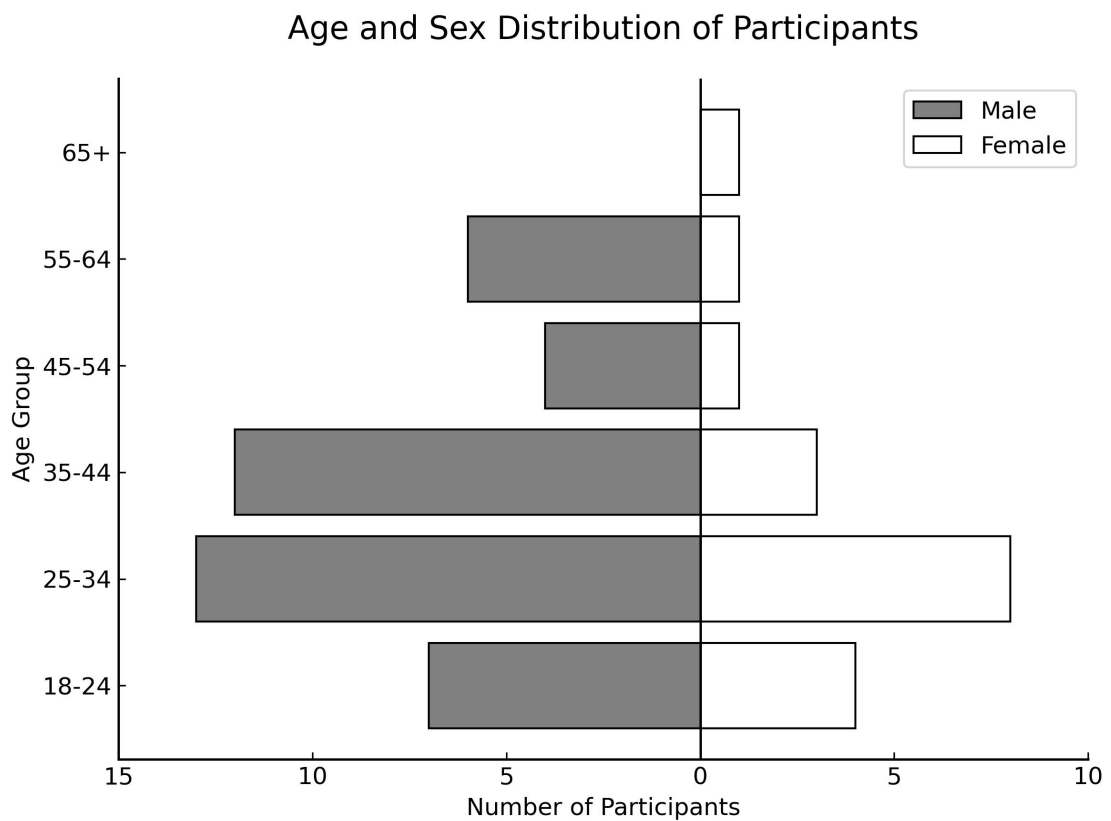


Figure 3.2 Demographic distribution of the final participant sample (N=60).

3.3 The experimental application

To conduct this experiment, a full-stack e-commerce web application was developed from the ground up. The application was designed to be a realistic simulation of a modern online store, complete with user authentication, product browsing, a shopping cart, and a multi-step checkout process.

3.3.1 Technology stack and architecture

The application is built on a modern JavaScript technology stack with a separated backend and frontend.

Backend — The server-side application is built with **Node.js** and the **Express.js** framework. It is responsible for all core business logic, including:

- Serving the frontend application’s static files.
- Providing a RESTful API for all data operations (e.g., fetching products, authenticating users).
- Managing the user database and all experimental data persistence.
- Handling the random assignment of participants to experimental groups.

The database is a single-file **SQLite** database, chosen for its simplicity and portability, making it ideal for this experimental setup. All participant interactions, metrics, and survey responses are stored in this database.

Two custom seeding scripts were developed to populate the database. The first script, `seed-users.js`, generates 200 unique username and password combinations, hashes the passwords using `bcrypt`, stores them in the database, and exports the plaintext credentials to a CSV file for distribution via Prolific. The second script, `seed-products.js`, populates the database with 60 fictional products and their variants, dynamically generating realistic data such as ratings, review counts, and stock levels. It also assigns dark pattern attributes (e.g., urgency text, countdown timers) to a random subset of products.

Frontend — The client-side application is a Single Page Application (SPA) built with **React.js**. The component-based architecture of React allowed for the creation of two distinct user experiences by swapping components based on the participant’s assigned group. Key libraries used include:

- **React router:** For handling client-side navigation between pages (e.g., Shop, Cart, Checkout).
- **React context API:** Used extensively to manage global state, including the user’s authentication status, shopping cart contents, experimental metrics, and, most importantly, the participant’s assigned version (A or B).
- **Axios:** For making asynchronous HTTP requests to the backend API.

Product imagery — To avoid stepping into the gray zone of copyright laws, all product images used in the application were generated using artificial intelligence. A custom Python script was developed, leveraging the **Stable Diffusion XL** model via the **diffusers** library. This script programmatically generated unique, high-resolution images for each of the 60 products and their visual variants based on descriptive prompts (e.g., "Product photo of [Product Name], [Variant Name], isolated on white background, 4k product render, hyperrealistic"). In cases where the initial automated generation produced suboptimal results, the images were manually curated or regenerated using various online AI image generation tools, including (among others) Google’s Imagen 3 & 4, OpenAI’s DALL-E and Microsoft Copilot, to achieve a consistent and believable aesthetic across the entire product catalogue.

Deployment — The entire full-stack application was deployed as a single service on **Render**, a modern cloud hosting platform. The server was configured to run on Render’s "Starter" plan, providing the necessary resources and importantly, a persistent disk. The persistent disk was essential for hosting the SQLite database file, so that all experimental data was preserved across deployments and server restarts. The frontend’s static build files were served directly by the Node.js backend, creating a unified and self-contained application accessible via a single public URL for the participants. Environment variables on Render were used to manage production-

specific configurations, such as the public-facing backend URL and secret keys for administrative access to download the database.

3.3.2 Version A: The dark pattern interface

Version A of the application was designed to implement the dark patterns defined in Chapter 2. The implementation focused on creating an environment of urgency and applying manipulative nudges throughout the user journey.

Visual design and atmosphere — The interface intentionally employed a color palette dominated by urgent reds and oranges, particularly for calls-to-action, warnings, and promotional banners. This was designed to amplify the psychological effects of the urgency and scarcity patterns. Gamified elements, such as progress bars for unlocking discounts or free shipping, were used to encourage users to increase their cart value (see Figure 3.7a). The specific implementations are detailed below, categorized by the pattern types from Section 2.1:

- **Misdirection:** Confirmshaming was implemented via modal windows that used emotive, guilt-inducing language. These modals were triggered whenever a user attempted an opt-out action, such as removing the pre-selected warranty. As seen in Figure 3.3, the modals also employed visual interference, styling the business-preferred option as a prominent button while de-emphasizing the user’s intended action.
- **Urgency and scarcity:** Fake countdown timers, high-visibility urgency tags (e.g., "Exclusive Offer"), and low-stock messages (e.g., "Only 8 left in stock!") were algorithmically added to product cards, as shown in Figure 3.4. This created a pervasive sense of time pressure and fear of missing out.
- **Pressured selling & social proof:** On the product detail modal, the default selected product variant was always the most expensive one. To justify this selection and leverage social proof, this variant was presented with artificially inflated metrics, showing significantly higher review counts and numbers of items sold compared to cheaper variants. Additionally, a social proof pop-up

(Figure 3.5) would periodically appear, claiming another user had just purchased the item.

- **Sneaking:** Two forms of sneaking were implemented. First, an optional product warranty was pre-selected by default for every item added to the shopping cart. Second, a mandatory "processing fee" was implemented as a hidden cost, only being revealed on the final checkout page.
- **Obstruction:** The product grid on the homepage was re-randomized on every page load, making it difficult for a user to navigate back to a previously seen item. Furthermore, the path to ending the experiment on the post-checkout page was obstructed by a final confirmshaming modal.

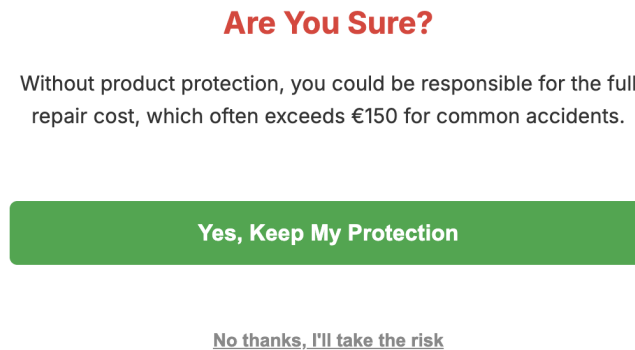


Figure 3.3 A Misdirection modal that combines a Confirmshaming message with Visual interference, where the business-preferred option ("Yes, Keep My Protection") is styled as the primary action.



Figure 3.4 A product card in Version A featuring multiple scarcity and urgency patterns: an "Exclusive Offer" tag, a low stock message, and a countdown timer.

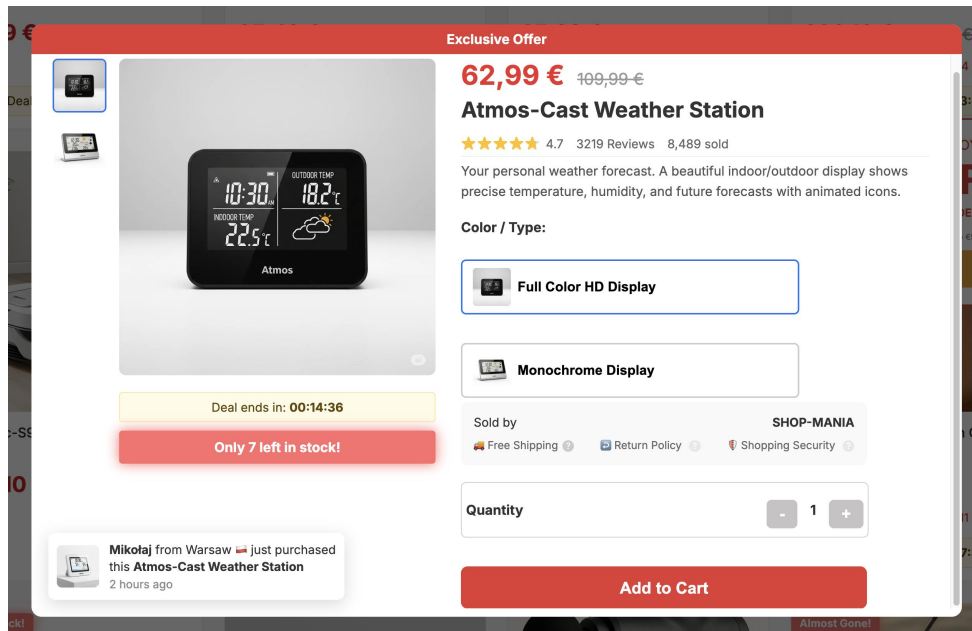


Figure 3.5 A social proof pop-up visible in the lower left corner of the product detail modal, designed to imply popularity and encourage a purchase.

3.3.3 Version B: The ethical design interface

Version B was designed as a direct counterpoint to Version A, implementing ethical alternatives for every dark pattern.

Visual design and atmosphere — The interface used a calm, neutral color palette dominated by blues and greys. All information was presented factually, without gamified or emotionally-charged elements. The specific ethical design choices were:

- **No false urgency:** All countdown timers, urgency banners, and low-stock warnings were removed entirely.
- **Honest defaults for products:** The default selected variant was always the cheapest, and social proof metrics were averaged across all variants. No fake social proofs pop ups were shown.

- **Transparent pricing:** All costs were disclosed upfront in the shopping cart, with tooltips for explanation.
- **Opt-in choices:** All optional add-ons were unchecked by default.
- **Frictionless navigation:** All confirmshaming modals were removed, the product grid was stable, and the post-checkout page presented clear, equally-weighted choices.

3.3.4 Side-by-side interface comparison

A detailed comparison of the key differences across the user journey is presented below, with explanatory text followed by the corresponding figures.

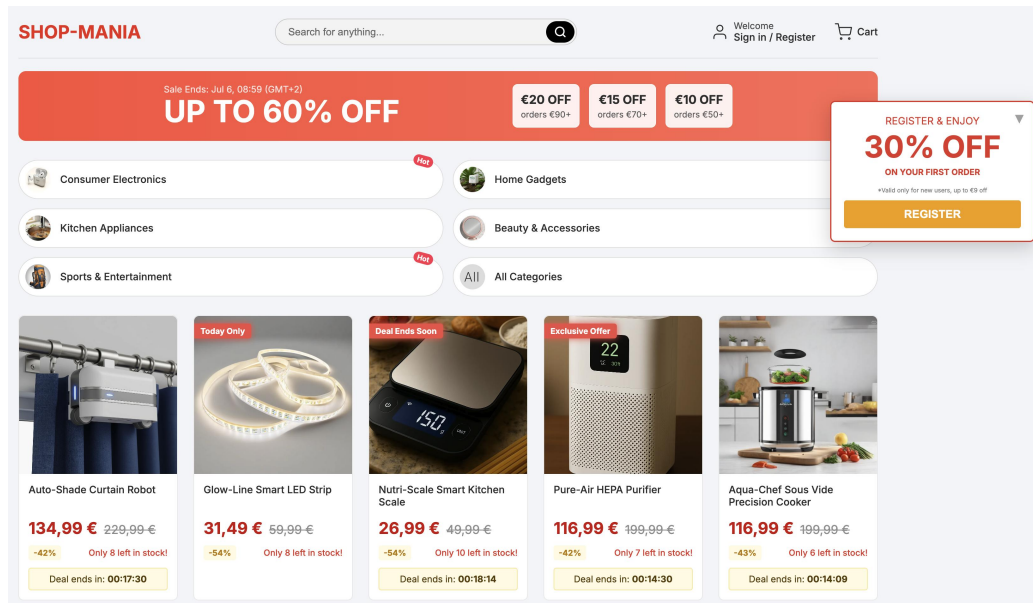
Home Page

The comparison in Figure 3.6 illustrates the starkly different browsing environments. Version A (3.6a) immediately presents a high-pressure atmosphere. Its red color scheme, multiple urgency tags on products (e.g., "Deal Ends Soon"), and the flashing promotional banner for registration all work together to create a sense of anxiety and fear of missing out. In contrast, Version B (3.6b) uses a calm blue palette. Discounts are shown as simple, factual percentages without any time pressure, creating a relaxed and trustworthy browsing experience. It employs no registration pop ups to nudge the user into registering, and provides a stable layout throughout the user experience.

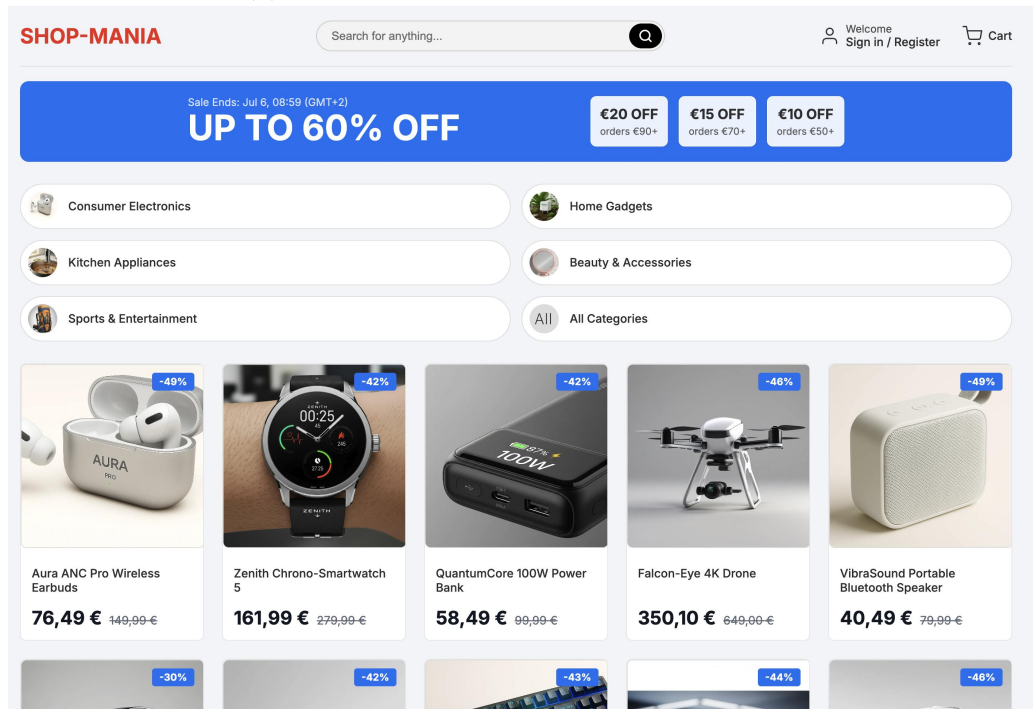
Shopping Cart

The shopping cart, shown in Figure 3.7, highlights the use of the Sneaking pattern. Version A (3.7a) pre-selects a warranty, adding it to the cart by default. Its order total is deceptive because the processing fee is still hidden. It also uses social proof ("9 other people have this in their cart") and a gamified progress bar to manipulate the user into spending more. Version B (3.7b) is transparent: the warranty is a clear opt-in choice, all fees are disclosed with explanatory tooltips, and the checkout options are presented neutrally.

Chapter 3. Methodology

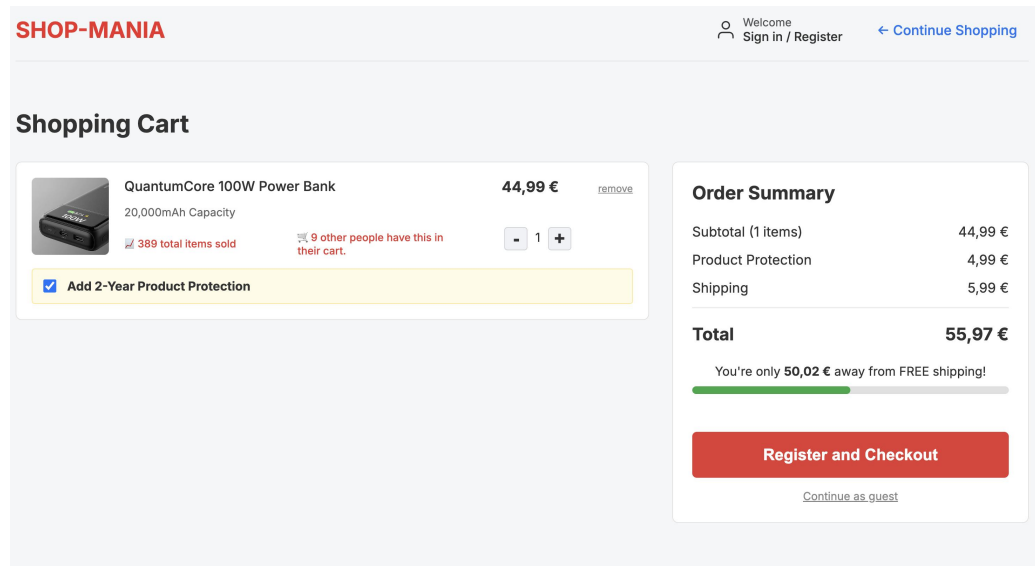


(a) Version A: High-pressure, urgent design.

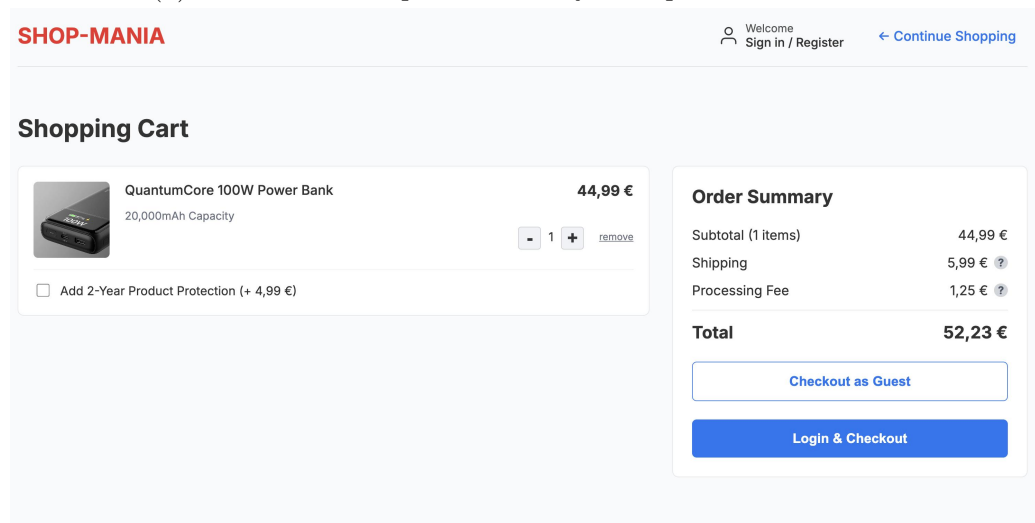


(b) Version B: Calm, transparent design.

Figure 3.6 Comparison of the home page.



(a) Version A: Deceptive summary with pre-selected items.



(b) Version B: Transparent summary with opt-in choices.

Figure 3.7 Comparison of the shopping cart.

Checkout Page

At the checkout stage, shown in Figure 3.8, Version A (3.8a) introduces the hidden processing fee for the first time, a common drip pricing tactic. It continues to apply pressure with a countdown timer and adds another pre-selected "Checkout Protection" fee. The card details are also pre-selected to be saved and in case of opting out a confirm shame modal would appear. The same goes if the user clicks the deemphasized return to cart link. Version B (3.8b) maintains consistency by showing the same total from the cart, with all choices remaining opt-in and no pressure elements or confirmshame modals being applied. The button to return to the cart is also more clearly presented.

Post-checkout journey

Finally, the post-checkout journey demonstrates another example of obstruction. As shown in Figure 3.9, in Version A users must navigate a two-step process where their intention to finish is interrupted by a confirmshaming modal. In contrast, Figure 3.10 shows the ethical design, where users are clearly presented two equal choices; to either continue or finish. Underneath, the "You Might Also Like" section contains static products unlike the animated carousel wheel in Version A.

← Return to cart

Shipping Information

John Smith

Avenue Louise

Belgium

+385912223344

Payment Details

Full name (as displayed on card)*
Card number*
Card expiration*
CVV*

Save card details for future purchases ☒

Your Order

Deal ends in: 00:04:40

Item Total	44.99 €
Shipping	5.99 €
Product Protection	4.99 €
Processing fee	1.25 €
Checkout Protection ☒	1.00 €

Enter coupon code

Add 45.01 € more to save 20 €!

Total 58.22 €

(a) Version A: Hidden costs revealed at the last minute.

← Return to Cart

Shipping Information

John Smith

Avenue Louise

Belgium

+385912223344

Payment Details

Full name (as displayed on card)*
Card number*
Card expiration*
CVV*

Save card details for future purchases ☐

Your Order

Item Total	46.24 €
Shipping	5.99 €

☐ Add Order Protection

Enter coupon code

Total 52.23 €

(b) Version B: Consistent and transparent pricing.

Figure 3.8 Comparison of the checkout page.

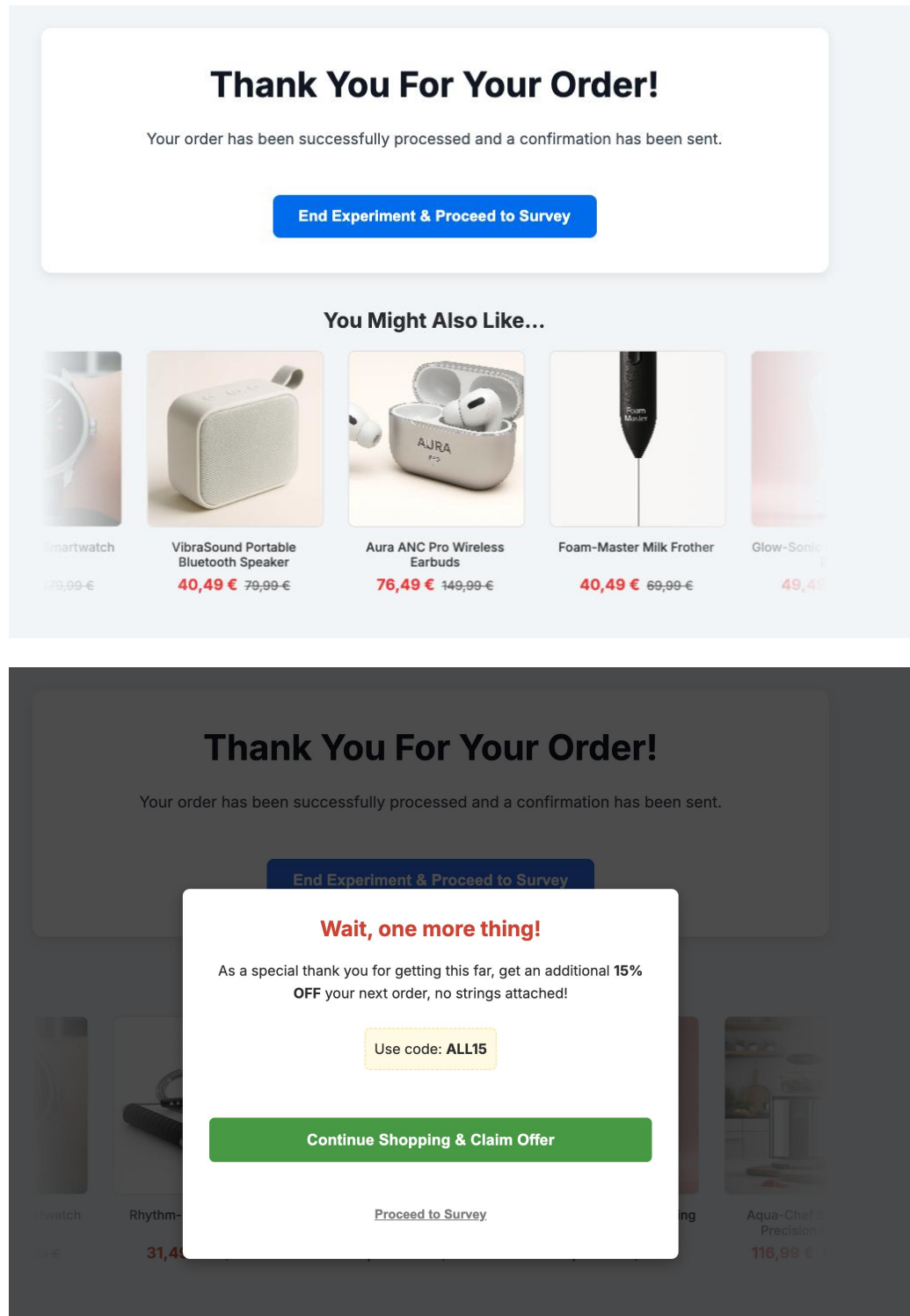


Figure 3.9 The two-step post-checkout journey in Version A, showing the confirmation modal.

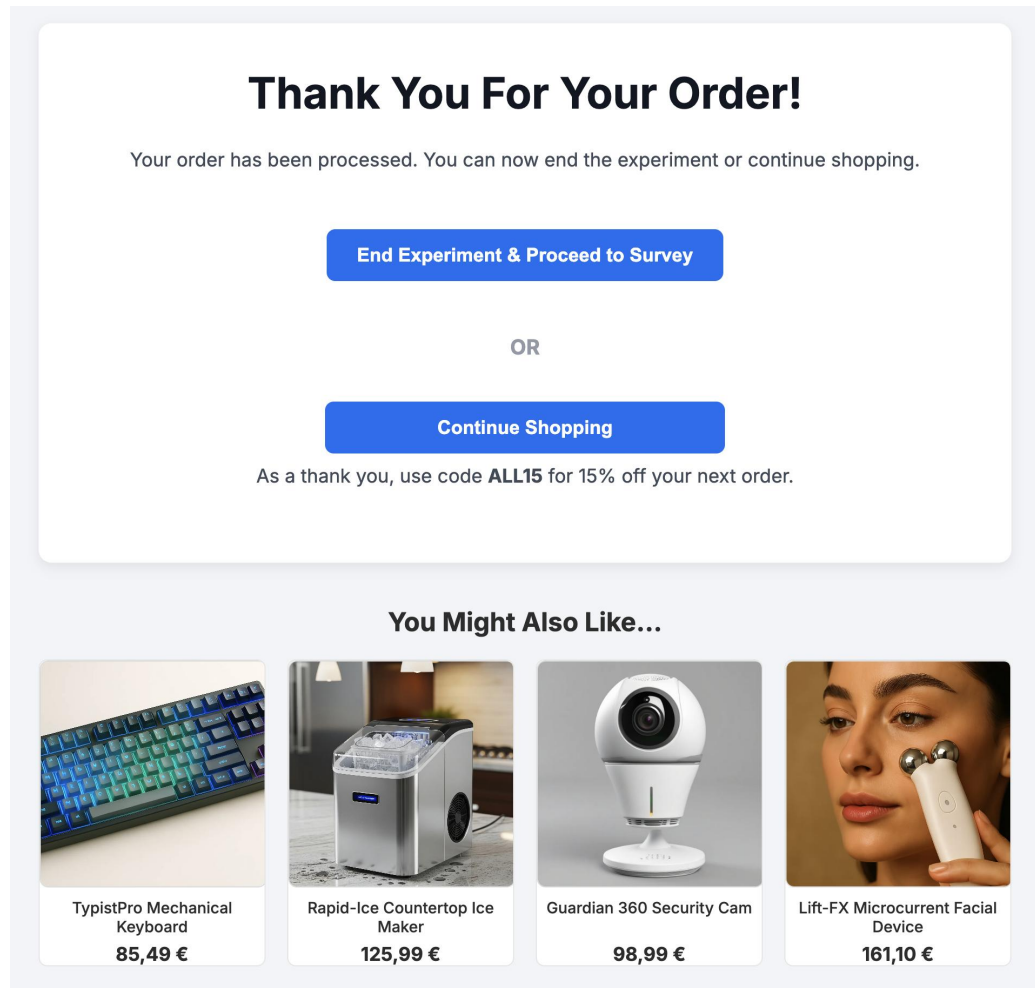


Figure 3.10 The single-step post-checkout journey in Version B.

3.4 Experimental procedure

The experiment was designed to guide participants through a complete and realistic online shopping journey, from initial onboarding to final survey submission. Interactions were constantly being monitored, updated and finally, saved to the database. The entire process was automated by the web application, so that every interaction was consistent for all participants.

3.4.1 Pre-task experience and onboarding

Initially, prospective participants were presented with the study description on the Prolific platform, as detailed in Section 3.2. This screen provided a neutral overview of the task, instructions on the simulated nature of the purchase and the estimated completion time. Upon accepting the study, participants were given their unique login credentials and a link to the "Shop-Mania" web application.

Once on the application's landing page, all participants, regardless of their future group assignment, underwent a standardized four-step onboarding process. This process was implemented in a way where every participant will have understood the task and possessed the necessary information to complete it. The steps were as follows:

1. **Introduction:** Participants were informed that their task was to simulate a shopping experience and purchase at least one item. They had to acknowledge this instruction to proceed.
2. **Fake credit card information:** A screen displayed the fake credit card details required for checkout (Card Number, Expiry Date, CVV). Participants were instructed to write down these details and confirm they had done so. In case they wrote down the wrong details in the checkout page, a pop up window would appear reminding them of the correct information.
3. **Rules & information:** Participants were explicitly told that no real money would be used, no personal information was required, and all products were fictional.

4. **Participant login:** The final step required participants to enter the unique username and password provided by Prolific (created by the `seed-users.js` script detailed in Section 3.3) to begin the experiment.

Figure 3.11 illustrates the key screens of this standardized onboarding procedure.

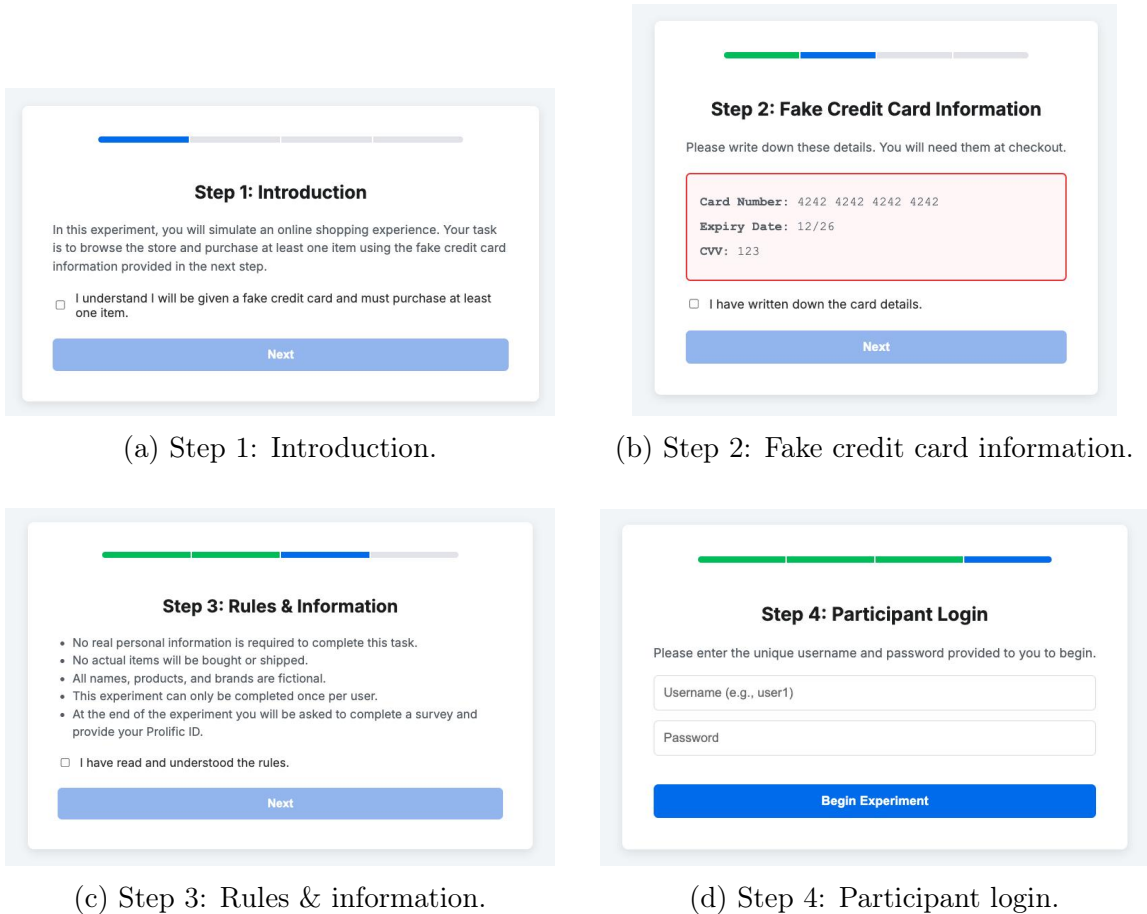


Figure 3.11 The four key screens of the standardized onboarding procedure.

3.4.2 Shopping experience and data collection

Upon successful login, the application's backend randomly assigned the participant to either Group A or Group B. This assignment was stored in the user's session, and from this point forward, their experiences diverged significantly.

Throughout the entire shopping task, the application's frontend, utilizing React's Context API, implemented as `MetricsContext` was responsible for logging events such as product impressions, clicks, time spent on different pages, and encounters with specific dark patterns. Simultaneously, the `CartContext` and `AuthContext` managed the state of the user's shopping cart and registration data, respectively. All collected data was stored in the browser's local storage and was only transmitted to the server upon final survey submission.

Journey of a Group A participant (Dark pattern)

A participant in Group A was directed to a homepage where the product grid was randomized on every page load, making re-discovery of items difficult. The interface was immediately populated with urgent red banners, countdown timers, and low-stock warnings on product cards. If the user was not registered, a flashing promotional banner for new users would appear after a short scroll, partially obstructing the view.

Upon clicking a product, the participant was shown a detail modal where the default selected variant was always the most expensive one, justified by artificially inflated review and sold counts. The sense of urgency was intensified with a countdown timer and a social proof pop-up claiming others had just purchased the item. When an item was added to the cart, a pre-selected product warranty was automatically included. On the cart page, this was highlighted in a yellow box to appear like a special offer. The order summary deliberately omitted the processing fee and used a gamified progress bar to nudge the user towards free shipping. If the participant attempted to deselect the warranty, a confirmshaming modal would appear, warning them of financial risk and using visual interference to make keeping the warranty the path of least resistance. To proceed, an unregistered user faced a large, red "Register and Checkout" button, while the "Continue as guest" option was a de-emphasized link. Clicking the guest link triggered another confirmshaming modal highlighting the disadvantages of not creating an account. The registration process itself was an intentionally high-friction, multi-step procedure designed to collect optional data, such as birthdays and interests, under the pretense of "personalizing the experience."

On the checkout page, the previously hidden processing fee was revealed for the first time, alongside another pre-selected "Checkout Protection" fee. A 5-minute countdown timer added further pressure. If the participant attempted to navigate back to the cart or deselect the "save card details" option, they were again met with confirmshaming modals.

Finally, after completing the purchase, the participant landed on the post-checkout page featuring an animated carousel of other products. The primary button to "End Experiment & Proceed to Survey" was obstructed by one final confirmshaming modal, which presented a "15% OFF" offer to entice the user to continue shopping instead of finishing the study.

Journey of a Group B participant (Ethical design)

A participant in Group B landed on a homepage with a clean, stable layout using a calm blue color scheme. The product grid, while randomized once for novelty, remained consistent on subsequent visits, allowing for easy re-discovery of items. No flashing banners or false urgency cues were present.

Upon clicking a product, the detail modal showed the cheapest variant selected by default. All social proof metrics like reviews and sold counts were averaged across all variants, presenting an honest, unified view of the product's popularity. There were no countdown timers or manipulative social proof pop-ups. When adding an item to the cart, no optional warranties were pre-selected. On the cart page, the warranty was presented as a clear, unchecked opt-in choice. The order summary was fully transparent, including the processing fee and estimated shipping costs from the outset. Helpful tooltips explained what each fee was for and how the user could qualify for free shipping. If an unregistered user chose to check out, they were presented with two equally weighted, clearly styled buttons: "Checkout as Guest" and "Login & Checkout." No shame or friction was associated with either choice. The registration process was straightforward and clearly marked optional fields as such, respecting user privacy. At checkout, there were no financial surprises. The total displayed was consistent with the total shown in the cart. The "Checkout

Protection" was unchecked and navigating back to the cart was a one-click action with no intervening modals. The box to save credit card details was opt-in instead of opt-out.

Finally, the post-checkout page thanked the user and presented a static grid of product recommendations. Two clear, equally weighted buttons offered the choice to "End Experiment & Proceed to Survey" or "Continue Shopping." Below the "Continue Shopping" button, a note explicitly offered a "15% OFF" coupon code as a reward, without forcing the user through a modal or obscuring the option to finish the study.

3.4.3 Post-experiment survey and completion

After successfully completing the checkout process, all participants, regardless of their group, were directed to the same post-experiment survey page. Here, they were asked to provide their Prolific ID needed for the demographics data, and answer six mandatory Likert-scale questions about their experience (see Section 3.1). Upon submitting the survey, a final API call was made to the `/api/metrics/save` endpoint. This single request transmitted the entire collection of data: the `user_metrics` JSON object, the `order_history`, `registration_data`, and the `survey_answers`, to the server where it was saved to the SQLite database. Once the data was successfully saved, the participant was redirected to the final completion page, which displayed a unique Prolific completion code. The participant would then copy this code and paste it into the Prolific interface to finalize their submission and receive payment.

3.5 Data analysis

All collected data was joined together and exported from the SQLite database into a CSV file for statistical analysis. The analysis was conducted programmatically using Python (version 3.11) and its scientific computing libraries, primarily **Pandas** for data manipulation, **SciPy** for statistical testing, and **Matplotlib/Seaborn** for data visualization. The significance level (alpha) for all statistical tests was set at $p < .05$.

3.5.1 Quantitative analysis

To compare the objective, measurable outcomes between Group A and Group B, a multi-step analysis was performed.

Outlier handling — For continuous variables susceptible to extreme values (e.g., spending and time metrics), statistical outliers were identified using the **1.5*interquartile range (IQR) rule** [12]. The IQR is the range between the first quartile (Q1, the 25th percentile) and the third quartile (Q3, the 75th percentile) of the data. Any data point falling below $Q1 - 1.5 \times IQR$ or above $Q3 + 1.5 \times IQR$ was flagged as an outlier and excluded to create a "cleaned" dataset for parametric testing.

Statistical testing — A two-pronged approach was used to ensure the robustness of the findings.

- An **independent-samples t-test** was conducted on the cleaned datasets to compare the means of the two groups. This test is appropriate for comparing two independent groups on a continuous dependent variable.
- To validate these findings and account for any non-normality in the original data, a non-parametric **Mann-Whitney U test** was performed on the full, uncleaned dataset. This test compares the medians of two groups and is not sensitive to outliers or the assumption of normal distribution [13].

For discrete and binary metrics (e.g., registration rate, number of products clicked),

where the 1.5*IQR rule is not applicable, the independent-samples t-test was conducted on the full dataset.

3.5.2 Qualitative (survey) analysis

The data from the post-experiment survey, which used a 5-point Likert scale, is ordinal in nature. Therefore, the most appropriate method for comparing the responses between the two independent groups was the non-parametric **Mann-Whitney U test** [13]. For each of the six survey questions, this test was used to determine if there was a statistically significant difference in the median scores reported by participants in Group A versus Group B.

3.5.3 Advanced and demographic analyses

To explore more nuanced relationships within the data, a series of correlational tests and subgroup analyses were conducted.

Correlational analysis — the relationship between two continuous or ordinal variables was assessed using two methods to ensure robustness, particularly when spending data was involved.

- A **Pearson product-moment correlation (r)** was used to measure the linear relationship between two continuous variables on datasets where statistical outliers had been removed. This was the primary test for correlations involving spending or time metrics on cleaned data.
- A non-parametric **Spearman’s rank-order correlation (rho)** was used as a robustness check on the full datasets with outliers. It was also the sole test used for any correlation involving ordinal data (i.e., any of the 5-point Likert scale survey scores), as it is the appropriate method for ranked data.

Subgroup analysis — to compare the means of demographic subgroups (e.g., males vs. females within an experimental group), the **independent-samples t-test** was used.

Chapter 4

Results

4.1 Quantitative findings

This section presents the quantitative findings from the experiment, comparing objective user behaviors between Group A (Dark Pattern) and Group B (Ethical Design). The results are organized by metric type, beginning with spending metrics, followed by time and engagement metrics, and concluding with discrete and binary outcomes. For each metric, descriptive statistics are provided in tables, and the results of the statistical comparisons are illustrated with corresponding charts.

4.1.1 Spending metrics

Final order value

One outlier was removed from each group (Group A: €1580.37, Group B: €7021.87). There was no significant difference between Group A ($M = 214.91 \pm 199.70$) and Group B ($M = 192.40 \pm 202.54$), $t(56) = 0.43$, $p > .05$. A Mann–Whitney U test on the full dataset confirmed no significant difference, $U = 498$, $p > .05$ (Fig. 4.1). The descriptive results can be seen in Table 4.1.

Chapter 4. Results

Table 4.1 Descriptive statistics for the Final Order Value, comparing Group A (Dark) and Group B (Ethical) across both the full and cleaned datasets.

	Full Dataset		Cleaned Dataset	
Statistic	Group A	Group B	Group A	Group B
N	29	31	28	30
Mean (€)	262.00	412.70	214.91	192.40
SD (€)	320.54	1242.67	199.70	202.54
Median (€)	140.23	125.24	121.73	125.24

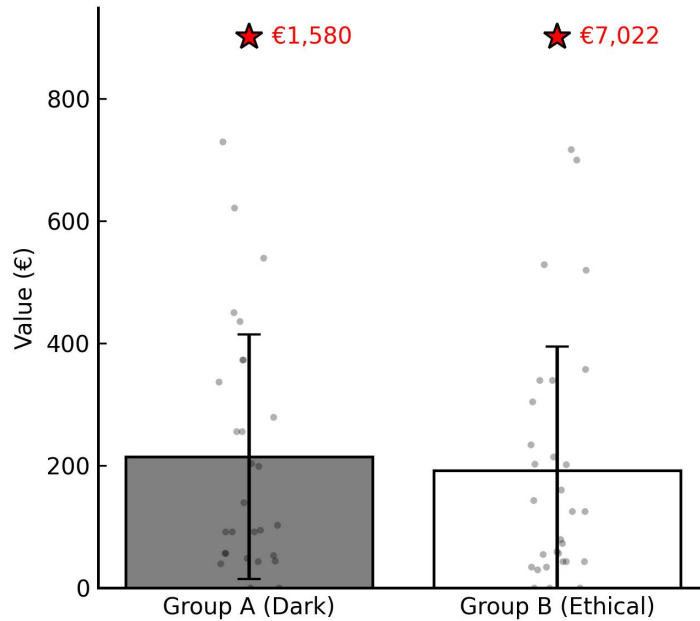


Figure 4.1 Comparison of Final order value (cleaned data).

Subtotal

One outlier was removed from each group (Group A: €1582.17, Group B: €6949.80). No significant difference was found between Group A ($M = 220.27 \pm 201.79$) and Group B ($M = 201.41 \pm 209.94$), $t(56) = 0.35$, $p > .05$. The Mann-Whitney U

Chapter 4. Results

test on the full dataset also indicated no significant difference, $U = 490$, $p > .05$ (Fig. 4.2). The descriptive results can be seen in Table 4.2.

Table 4.2 Descriptive statistics for the Subtotal, comparing Group A and Group B across both the full and cleaned datasets.

	Full Dataset		Cleaned Dataset	
Statistic	Group A	Group B	Group A	Group B
N	29	31	28	30
Mean (€)	267.23	419.10	220.27	201.41
SD (€)	321.28	1229.50	201.79	209.94
Median (€)	152.99	143.99	134.99	143.99

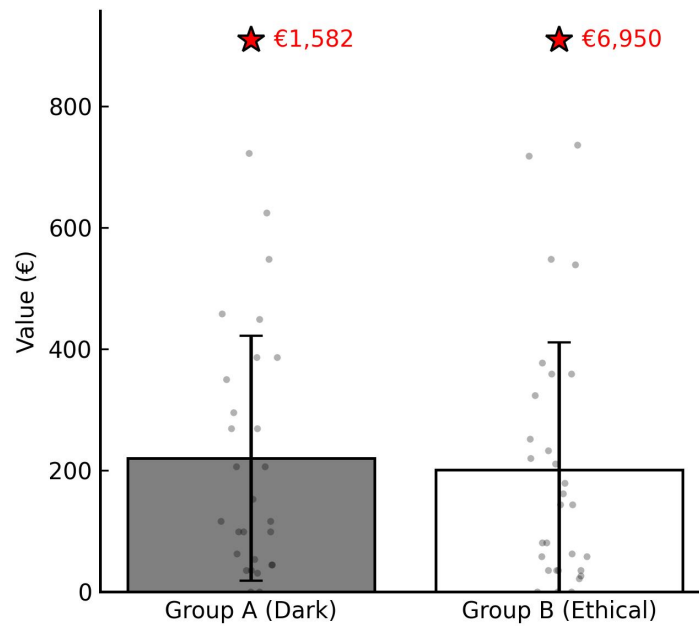


Figure 4.2 Comparison of Subtotal (cleaned data).

Cost discrepancy

One outlier was removed from Group B (€72.07). There was no significant difference between Group A ($M = -5.23 \pm 8.63$) and Group B ($M = -9.01 \pm 10.48$), $t(57) = 1.51$, $p > .05$. A Mann–Whitney U test on the full dataset confirmed no significant difference, $U = 542$, $p > .05$ (Fig. 4.3). The descriptive results can be seen in Table 4.3.

Table 4.3 Descriptive statistics for the Cost Discrepancy, comparing Group A and Group B across both the full and cleaned datasets.

	Full Dataset		Cleaned Dataset	
Statistic	Group A	Group B	Group A	Group B
N	29	31	29	30
Mean (€)	-5.23	-6.40	-5.23	-9.01
SD (€)	8.63	17.84	8.63	10.48
Median (€)	-6.77	-8.77	-6.77	-13.26

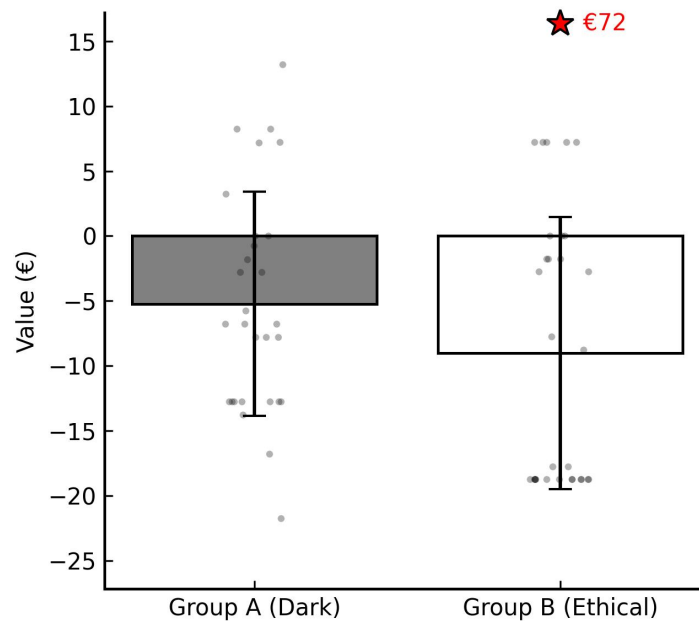


Figure 4.3 Comparison of Cost discrepancy (cleaned data).

4.1.2 Time & engagement metrics

Total session time

One outlier was removed from Group B (1046.16 s). Group A ($M = 287.94 \pm 77.73$ s) had a significantly longer session time than Group B ($M = 235.06 \pm 82.81$ s), $t(57) = 2.53$, $p = .014$. The Mann–Whitney U test on the full dataset also found a significant difference, $U = 600$, $p = .026$ (Fig. 4.4). The descriptive results can be seen in Table 4.4.

Table 4.4 Descriptive statistics for Total session time, comparing Group A and Group B across both the full and cleaned datasets.

	Full Dataset		Cleaned Dataset	
Statistic	Group A	Group B	Group A	Group B
N	29	31	29	30
Mean (s)	287.94	261.23	287.94	235.06
SD (s)	77.73	166.88	77.73	82.81
Median (s)	278.59	231.23	278.59	229.49

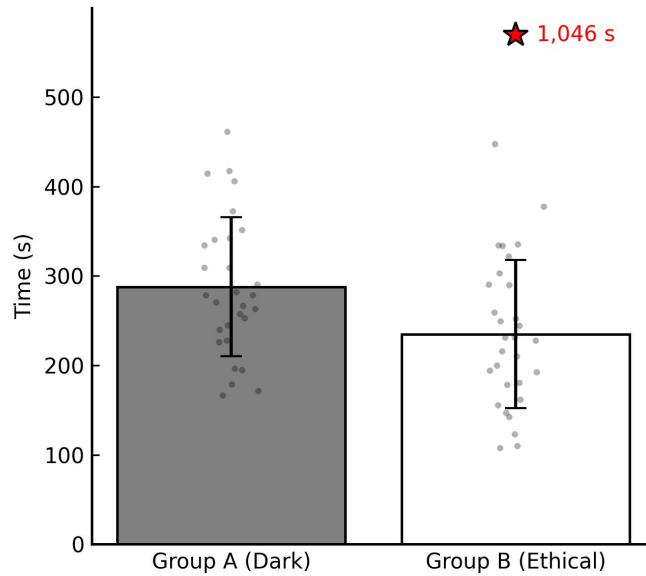


Figure 4.4 Comparison of Total session time (cleaned data).

Time to first engagement

One outlier was removed from Group A (246 s) and four from Group B (98 s, 197 s, 192 s, 109 s). Group A ($M = 46.12 \pm 31.68$ s) had a significantly longer time to first engagement than Group B ($M = 26.87 \pm 15.70$ s), $t(53) = 2.87$, $p = .007$. However, the Mann–Whitney U test on the full dataset did not find a significant difference, $U = 560$, $p > .05$ (Fig. 4.5). The descriptive results can be seen in Table 4.5.

Table 4.5 Descriptive statistics for Time to first engagement, comparing Group A and Group B across both the full and cleaned datasets.

	Full Dataset		Cleaned Dataset	
Statistic	Group A	Group B	Group A	Group B
N	29	31	28	27
Mean (s)	53.01	42.64	46.12	26.87
SD (s)	48.42	47.15	31.68	15.70
Median (s)	37.17	29.11	36.92	25.41

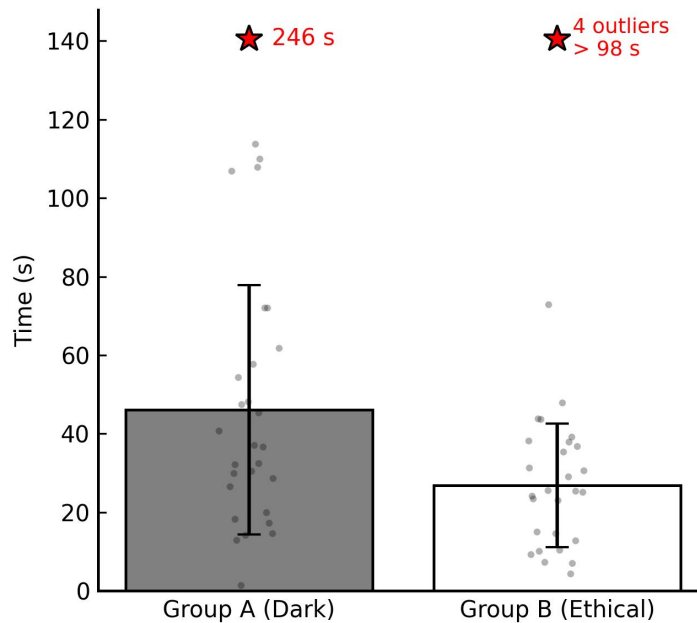


Figure 4.5 Comparison of Time to first engagement (cleaned data).

Passive browsing time

Three outliers were removed from Group A (175 s, 196 s, 275 s) and one from Group B (199 s). No significant difference was found between Group A ($M = 61.08 \pm 24.90$ s) and Group B ($M = 50.53 \pm 31.44$ s), $t(54) = 1.40$, $p > .05$. The Mann–Whitney U test confirmed no significant difference, $U = 580$, $p > .05$ (Fig. 4.6). The descriptive results can be seen in Table 4.6.

Table 4.6 Descriptive statistics for Passive browsing time, comparing Group A and Group B across both the full and cleaned datasets.

	Full Dataset		Cleaned Dataset	
Statistic	Group A	Group B	Group A	Group B
N	29	31	26	30
Mean (s)	77.03	55.32	61.08	50.53
SD (s)	55.12	40.82	24.90	31.44
Median (s)	63.00	49.00	59.50	46.50

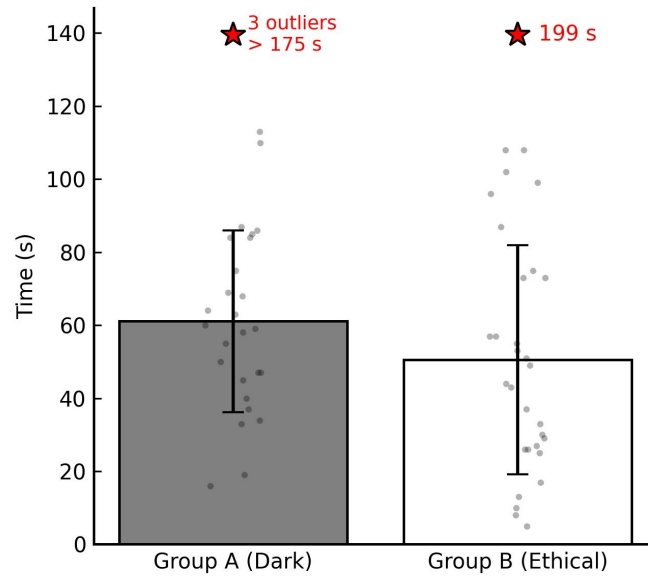


Figure 4.6 Comparison of Passive browsing time (cleaned data).

4.1.3 Discrete & binary metrics

Passively viewed products

Group A viewed significantly more products passively ($M = 39.38 \pm 16.19$) than Group B ($M = 29.48 \pm 17.17$), $t(58) = 2.30$, $p = .025$ (Fig. 4.7). The descriptive statistics are shown in Table 4.7.

Table 4.7 Descriptive statistics for Passively viewed products.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	39.38	29.48
SD	16.19	17.17
Median	40.00	25.00

Products clicked

There was no significant difference between Group A ($M = 1.93 \pm 1.31$) and Group B ($M = 1.94 \pm 1.41$) in products clicked, $t(58) = -0.01$, $p > .05$ (Fig. 4.8). The descriptive statistics are shown in Table 4.8.

Table 4.8 Descriptive statistics for Products clicked.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	1.93	1.94
SD	1.31	1.41
Median	1.00	1.00

Chapter 4. Results

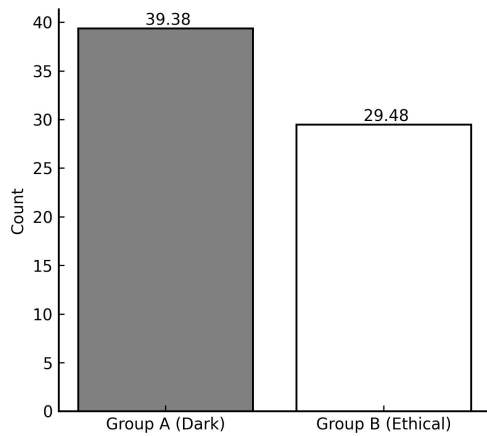


Figure 4.7 Passively viewed products.

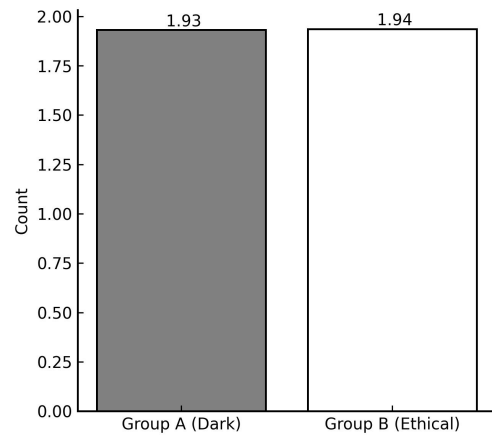


Figure 4.8 Products clicked.

Registration rate

Group A had a significantly higher registration rate ($M = 0.52 \pm 0.51$) than Group B ($M = 0.10 \pm 0.30$), $t(58) = 3.87$, $p < .001$ (Fig. 4.9). The descriptive statistics are shown in Table 4.9.

Table 4.9 Descriptive statistics for Registration rate.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	0.52	0.10
SD	0.51	0.30
Median	1.00	0.00

Checkout protection rate

Group A had a significantly higher checkout protection rate ($M = 0.93 \pm 0.26$) than Group B ($M = 0.29 \pm 0.46$), $t(58) = 6.69$, $p < .001$ (Fig. 4.10). The descriptive statistics are shown in Table 4.10.

Table 4.10 Descriptive statistics for Checkout protection rate.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	0.93	0.29
SD	0.26	0.46
Median	1.00	0.00

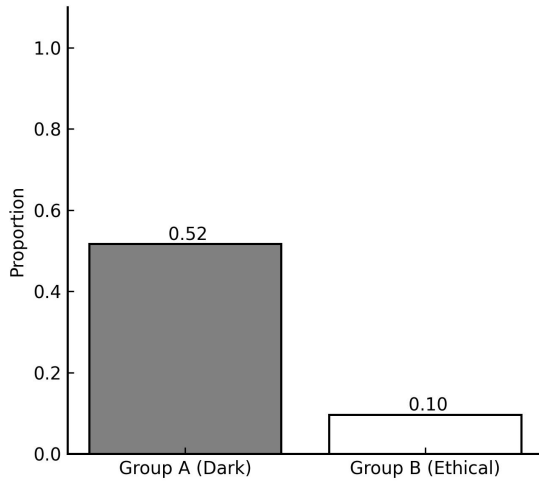


Figure 4.9 Registration rate.

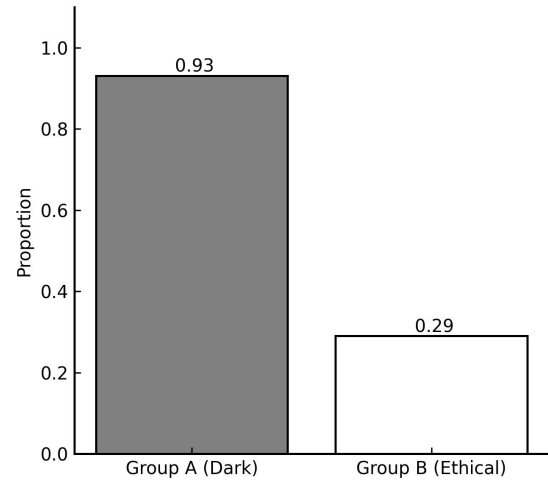


Figure 4.10 Checkout protection rate.

Warranty purchase rate

Warranty purchase rate was significantly higher in Group A ($M = 0.79 \pm 0.41$) than Group B ($M = 0.10 \pm 0.30$), $t(58) = 7.43$, $p < .001$ (Fig. 4.11). The descriptive statistics are shown in Table 4.11.

Table 4.11 Descriptive statistics for Warranty purchase rate.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	0.79	0.10
SD	0.41	0.30
Median	1.00	0.00

Saved card rate

Group A's saved card rate ($M = 0.79 \pm 0.41$) was significantly higher than Group B's ($M = 0.10 \pm 0.30$), $t(58) = 7.43$, $p < .001$ (Fig. 4.12). The descriptive statistics are shown in Table 4.12.

Table 4.12 Descriptive statistics for Saved card rate.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	0.79	0.10
SD	0.41	0.30
Median	1.00	0.00

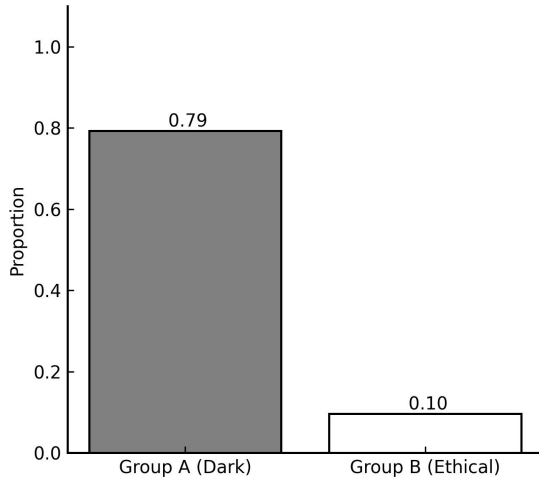


Figure 4.11 Warrantly purchase rate.

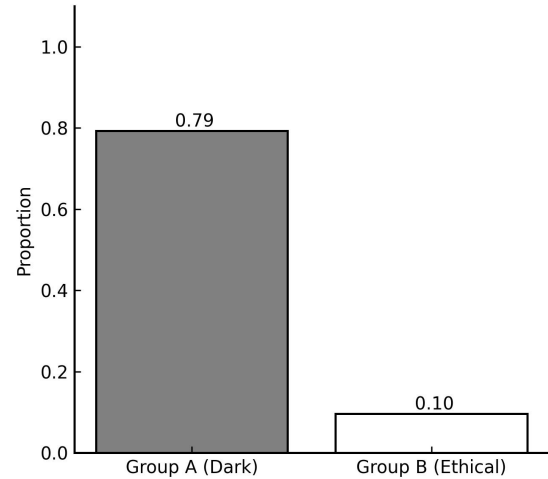


Figure 4.12 Saved card rate.

Provided birthday rate

Group A had a significantly higher provided birthday rate ($M = 0.48 \pm 0.51$) than Group B ($M = 0.06 \pm 0.25$), $t(58) = 4.00$, $p < .001$ (Fig. 4.13). The descriptive statistics are shown in Table 4.13.

Table 4.13 Descriptive statistics for Provided birthday rate.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	0.48	0.06
SD	0.51	0.25
Median	0.00	0.00

Provided interests rate

Group A's provided interests rate ($M = 0.34 \pm 0.48$) was significantly higher than Group B's ($M = 0.03 \pm 0.18$), $t(58) = 3.27$, $p = .002$ (Fig. 4.14). The descriptive statistics are shown in Table 4.14.

Table 4.14 Descriptive statistics for Provided interests rate.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	0.34	0.03
SD	0.48	0.18
Median	0.00	0.00

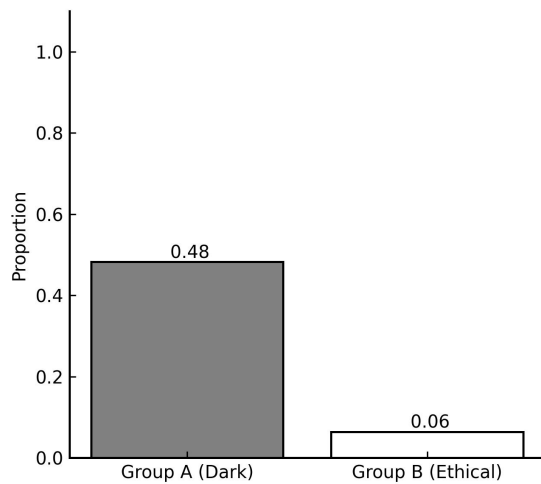


Figure 4.13 Provided birthday rate.

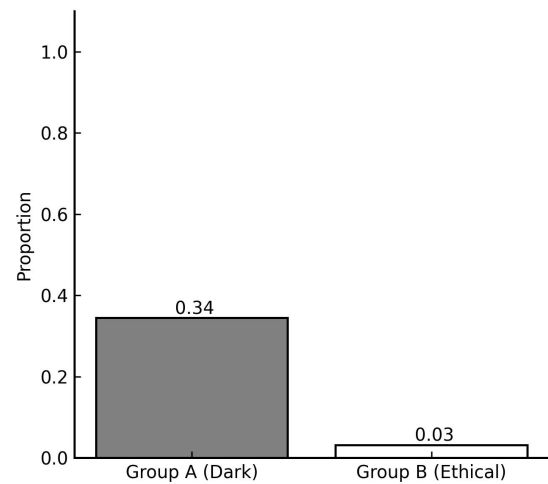


Figure 4.14 Provided interests rate.

4.2 Qualitative findings

This section presents the qualitative findings based on participant ratings for perceived ease and clarity of use, perceived control and trust, and awareness of manipulative elements. Each subsection reports descriptive statistics for Group A and Group B, followed by the results of the Mann–Whitney U test.

4.2.1 Perceived ease & clarity of use

The website was easy to navigate and understand (Q1)

Median scores for the perceived ease of use in Group A (Mdn = 5.0) and Group B (Mdn = 5.0) did not differ significantly, $U = 460$, $p > .05$ (Fig. 4.15). The descriptive statistics are presented in Table 4.15.

Table 4.15 Descriptive statistics for Q1: Website easy to navigate and understand.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	4.38	4.35
SD	0.78	0.75
Median	5.00	5.00

I always felt clear about the total cost of my order (Q2)

Median scores for the clarity in Group A (Mdn = 4.0) and Group B (Mdn = 4.0) did not differ significantly, $U = 396$, $p > .05$ (Fig. 4.15). The descriptive statistics are presented in Table 4.16.

Table 4.16 Descriptive statistics for Q2: Clarity about total cost.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	3.79	4.10
SD	1.21	0.94
Median	4.00	4.00

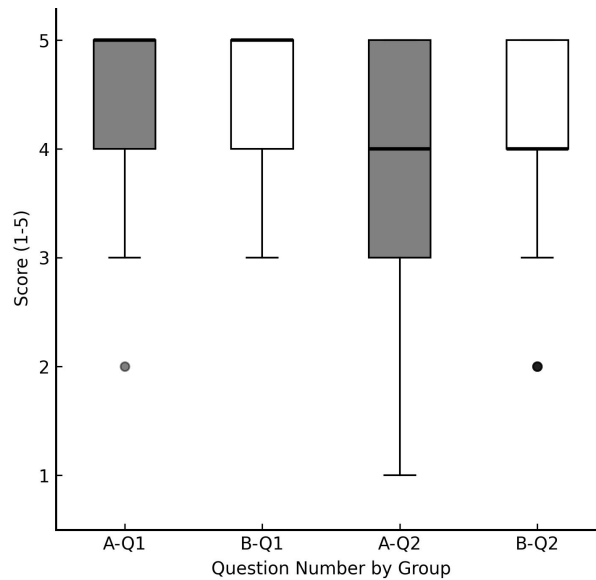


Figure 4.15 Perceived ease & clarity of use scores (Q1 & Q2).

4.2.2 Perceived control & trust

I felt in complete control of my decisions while shopping (Q3)

Median scores for perceived control in Group A (Mdn = 4.0) and Group B (Mdn = 4.0) did not differ significantly, $U = 438$, $p > .05$ (Fig. 4.16). The descriptive statistics are shown in Table 4.17.

Table 4.17 Descriptive statistics for Q3: Feeling in control of decisions.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	3.97	4.19
SD	1.21	0.70
Median	4.00	4.00

I trust this website to be transparent and act in my best interest (Q4)

Median scores for trust in Group A (Mdn = 3.0) and Group B (Mdn = 4.0) did not differ significantly, $U = 361$, $p > .05$ (Fig. 4.16). The descriptive statistics are shown in Table 4.18.

Table 4.18 Descriptive statistics for Q4: Trust in website transparency.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	3.17	3.61
SD	1.14	0.99
Median	3.00	4.00

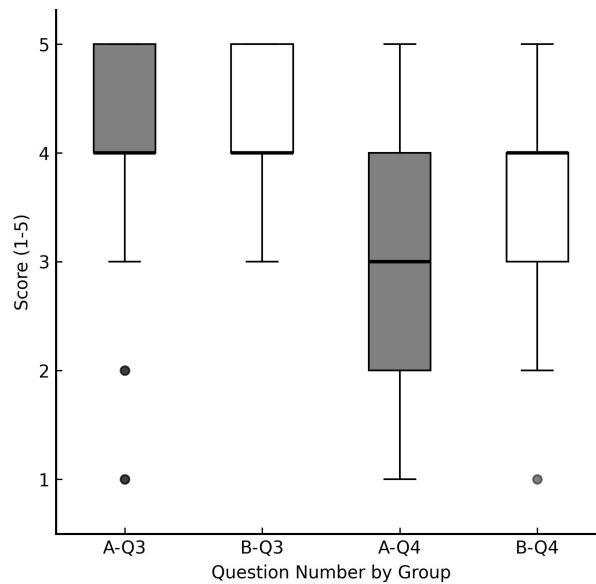


Figure 4.16 Perceived control & trust scores (Q3 & Q4).

4.2.3 Awareness of manipulative elements

I felt pressured to make a purchase or add extra items (Q5)

Median scores for feelings of pressure in Group A (Mdn = 2.0) and Group B (Mdn = 2.0) did not differ significantly, $U = 526$, $p > .05$ (Fig. 4.17). The descriptive statistics are shown in Table 4.19.

Table 4.19 Descriptive statistics for Q5: Feeling pressured to purchase.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	2.79	2.39
SD	1.42	1.28
Median	2.00	2.00

I noticed design elements that seemed intended to influence my purchasing decisions (Q6)

Group A (Mdn = 4.0) reported significantly higher awareness than Group B (Mdn = 3.0), $U = 606$, $p = .018$ (Fig. 4.17). The descriptive statistics are shown in Table 4.20.

Table 4.20 Descriptive statistics for Q6: Awareness of manipulative design elements.

Statistic	Group A (Dark)	Group B (Ethical)
N	29	31
Mean	3.69	2.94
SD	1.07	1.24
Median	4.00	3.00

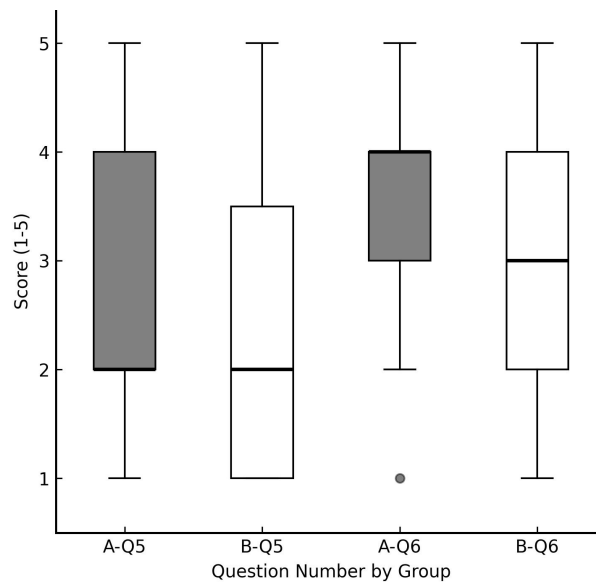


Figure 4.17 Awareness of manipulative elements scores (Q5 & Q6).

4.3 Advanced & demographic analyses

This section presents a deeper investigation into the factors that may influence user behavior and perception, moving beyond the primary comparison of the two experimental groups. The analysis explores the impact of a user's self-reported awareness on their spending, and the role of key demographics in relation to both spending habits and subjective survey responses.

4.3.1 Impact of awareness on behavior (Group A)

A Pearson product-moment correlation on the cleaned data found a weak, positive correlation between awareness score and final order value, which was not statistically significant ($r = 0.010$, $N = 28$, $p > .05$). A Spearman's Rho correlation on the full dataset confirmed this finding, showing a weak, non-significant negative correlation ($\rho = -0.056$, $N = 29$, $p > .05$). The relationship is visualized in Figure 4.18. The descriptive statistics are presented in Table 4.21.

Table 4.21 Descriptive statistics for Awareness analysis (Group A).

	Full Dataset		Cleaned Dataset	
Statistic	Score(1-5)	Final Total (€)	Score(1-5)	Final Total (€)
N	29	29	28	28
Mean	3.69	262.00	3.68	214.91
Std. Dev.	1.07	320.54	1.09	199.70
Median	4.00	140.23	4.00	121.73

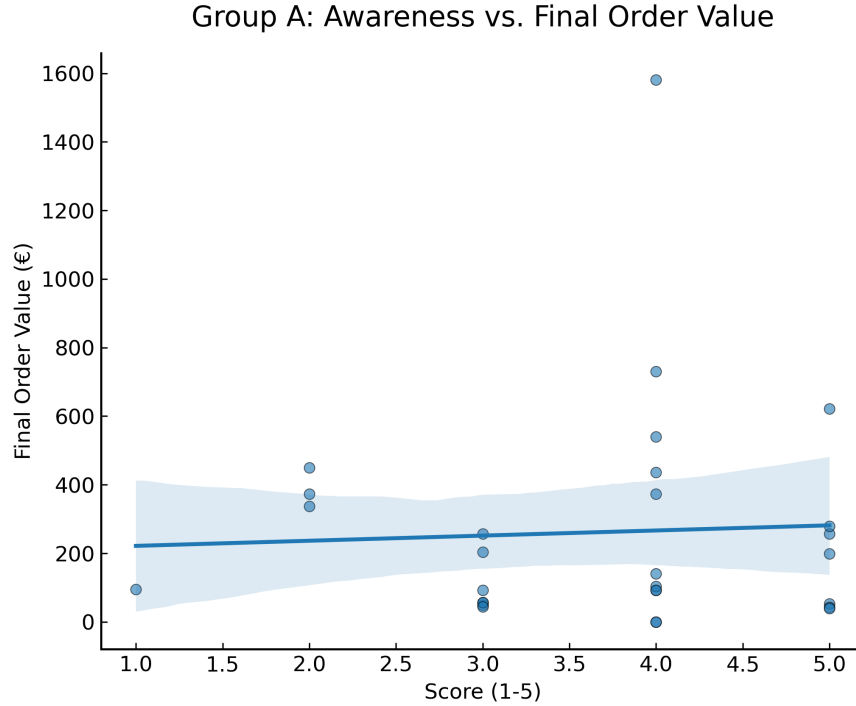


Figure 4.18 Correlation between Awareness score and Final order value (full dataset).

4.3.2 Demographic correlation and segmentation

This section explores the relationships between user demographics (age and sex), their spending behavior and subjective survey responses.

Correlation between Age and Final order value

A Pearson product-moment correlation was run on the cleaned data. In Group A, a weak, negative correlation was found, which was not statistically significant ($r = -0.043$, $N = 28$, $p > .05$). Similarly, in Group B, a weak, negative correlation was found, which was also not statistically significant ($r = -0.124$, $N = 30$, $p > .05$). Robustness checks using Spearman's Rho on the full dataset yielded similar non-significant results for both Group A ($\rho = 0.127$, $N = 29$, $p > .05$) and Group B ($\rho = -0.153$, $N = 31$, $p > .05$). The relationship is visualized in Figure 4.19. The

descriptive statistics are presented in Table 4.22.

Table 4.22 Descriptive statistics for Age and Final order value.

	Group A		Group B	
Statistic	Full Data	Cleaned Data	Full Data	Cleaned Data
N	29	28	31	30
Mean (€)	262.00	214.91	412.70	192.40
Std. Dev. (€)	320.54	199.70	1242.67	202.54
Median (€)	140.23	121.73	125.24	125.24

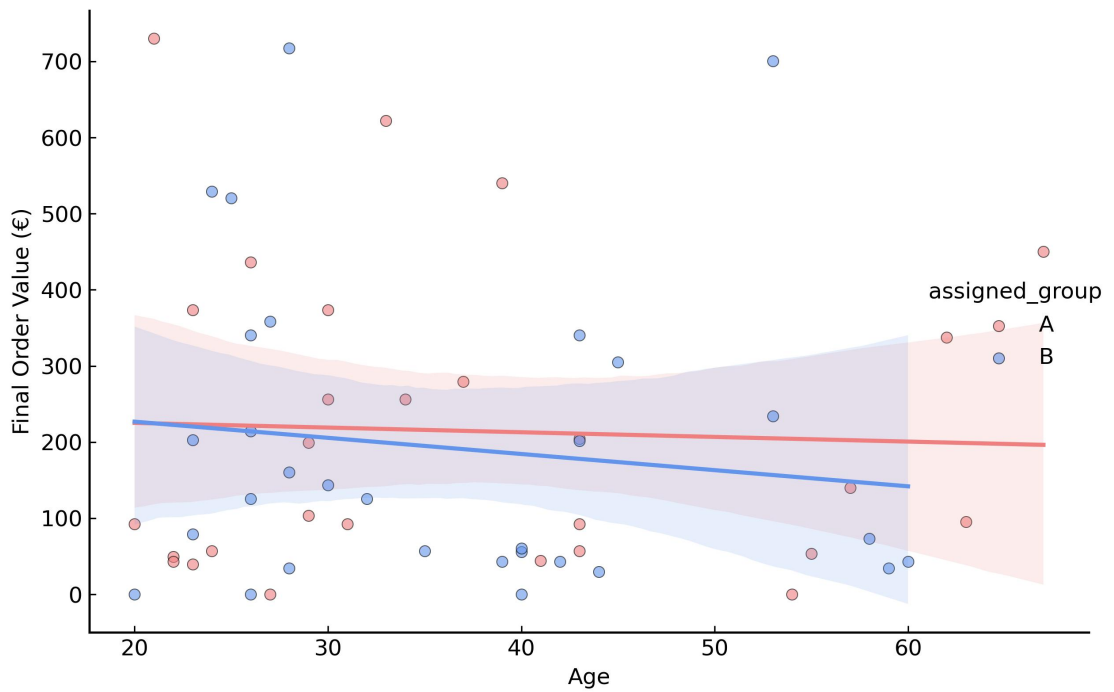


Figure 4.19 Correlation between Age and Final order value (cleaned data).

Correlation between Age and Subjective perceptions

Spearman's Rho correlations were run on the full dataset to determine the relationship between age and survey responses for the following subjective metrics:

Perceived control (q3) — Descriptive statistics for the Perceived Control scores are presented in Table 4.23. A statistically significant, moderate, positive correlation was found in Group A ($\rho = 0.605$, $N = 29$, $p < .001$) and in Group B ($\rho = 0.385$, $N = 31$, $p < .05$). In both groups, as age increased, the feeling of control increased. The relationship is visualized in Figure 4.20

Table 4.23 Descriptive statistics for Age and Perceived control (q3).

Statistic	Score (Group A)	Score (Group B)
N (Count)	29	31
Mean	3.97	4.19
Median	4.00	4.00

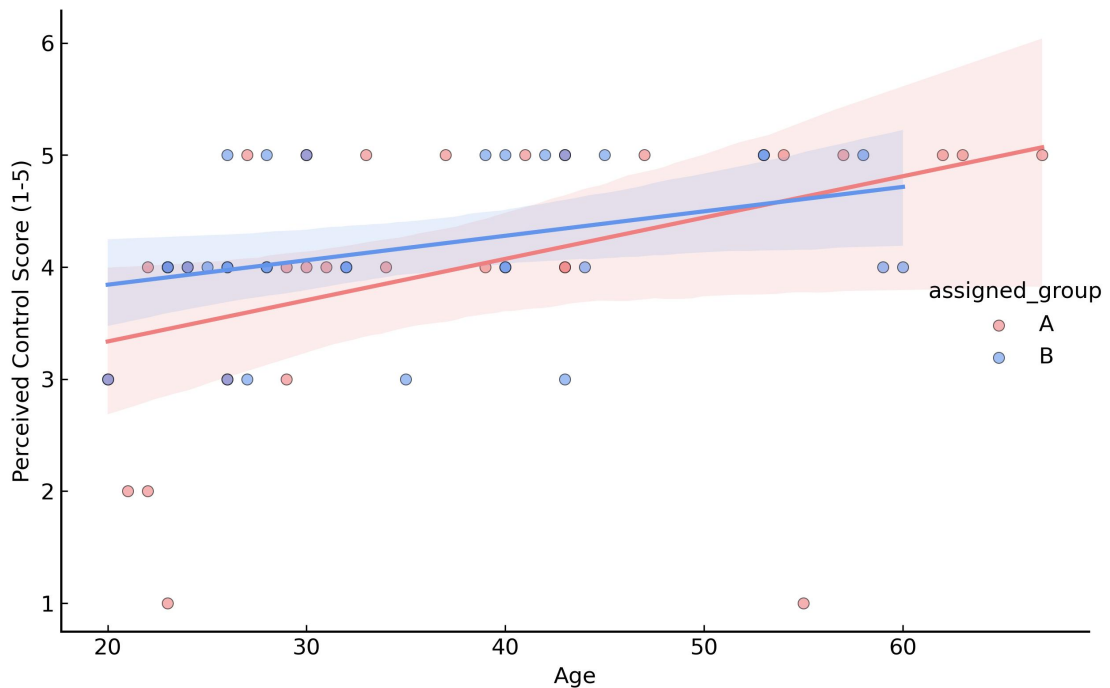


Figure 4.20 Correlation between Age and Perceived control score (q3).

Website Trust (q4) — Descriptive statistics for the Website trust scores are presented in Table 4.24. In Group A, a statistically significant, moderate, positive correlation was found ($\rho = 0.408$, $N = 29$, $p < .05$) as age increased website trust increased. No significant correlation was found in Group B ($\rho = -0.110$, $N = 31$, $p > .05$). The relationship is visualized in Figure 4.21

Table 4.24 Descriptive statistics for Age and Website trust (q4).

Statistic	Score (Group A)	Score (Group B)
N (Count)	29	31
Mean	3.17	3.61
Median	3.00	4.00

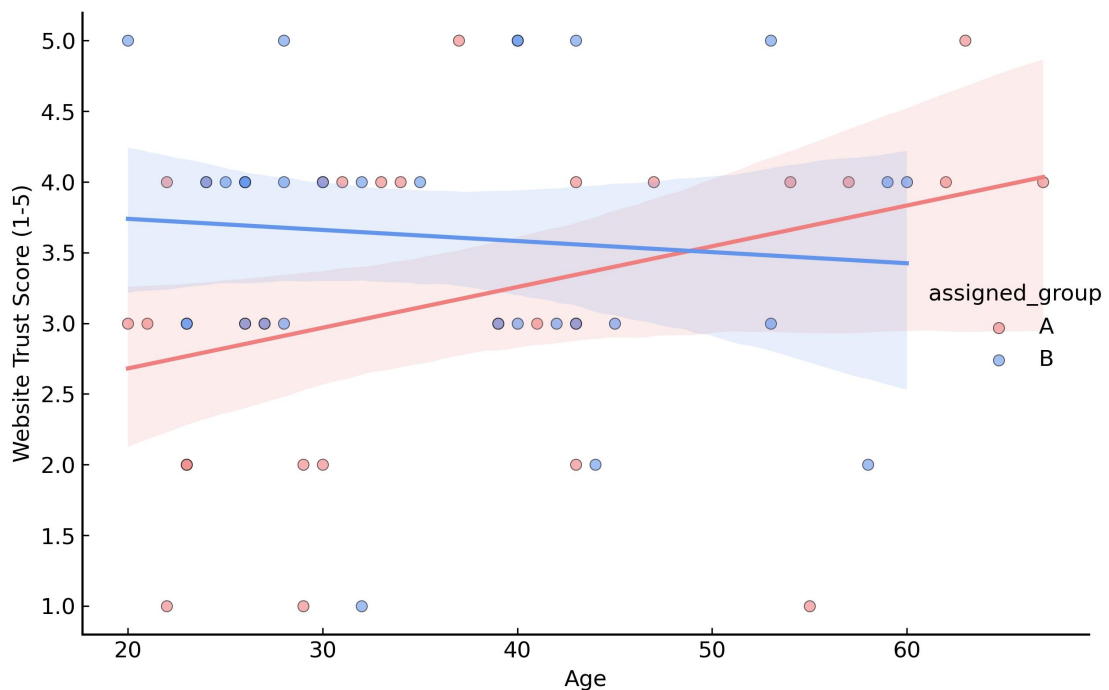


Figure 4.21 Correlation between Age and Website trust score (q4).

Feeling of pressure (q5) — Descriptive statistics for the Feeling of pressure scores are presented in Table 4.25. No statistically significant correlation was found

Chapter 4. Results

between age and this metric in either Group A ($\rho = -0.005$, $N = 29$, $p > .05$) or Group B ($\rho = -0.315$, $N = 31$, $p > .05$). The relationship is visualized in Figure 4.22

Table 4.25 Descriptive statistics for Age and Feeling of pressure (q5).

Statistic	Score (Group A)	Score (Group B)
N (Count)	29	31
Mean	2.79	2.39
Median	2.00	2.00

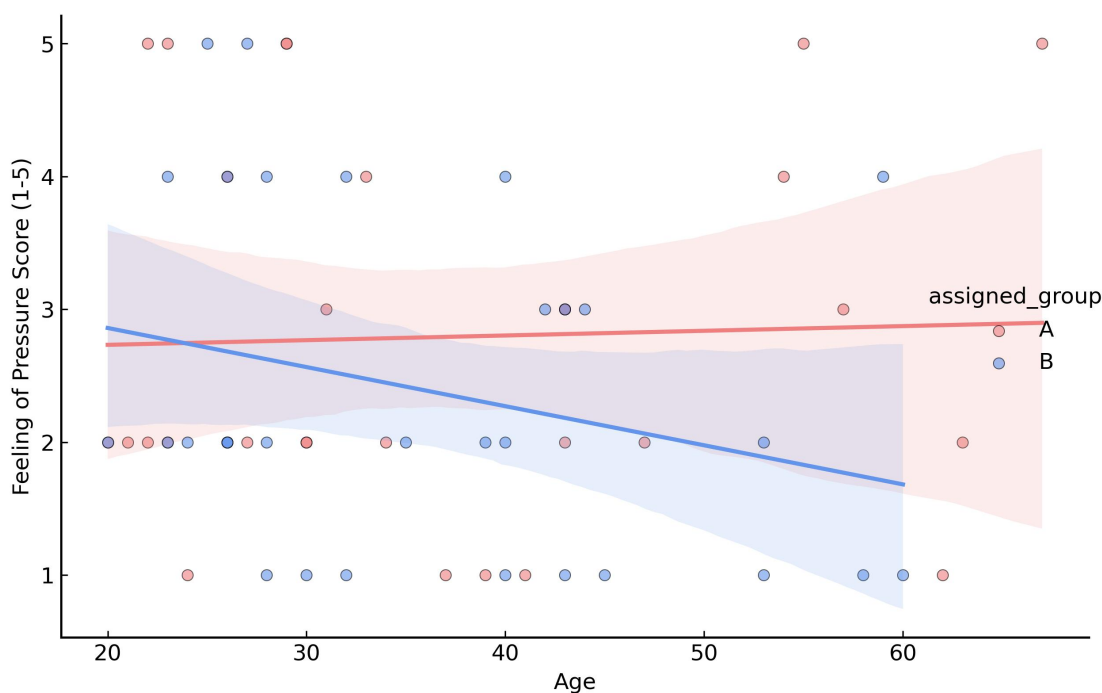


Figure 4.22 Correlation between Age and Feeling of pressure score (q5).

Awareness of manipulation (q6) — Descriptive statistics for the Awareness of Manipulation scores are presented in Table 4.26. No statistically significant correlation was found between age and this metric in either Group A ($\rho = -0.274$, $N = 29$, $p > .05$) or Group B ($\rho = -0.119$, $N = 31$, $p > .05$). The relationship is visualized in Figure 4.23

Table 4.26 Descriptive statistics for Age and Awareness of manipulation (q6).

Statistic	Score (Group A)	Score (Group B)
N (Count)	29	31
Mean	3.69	2.94
Median	4.00	3.00

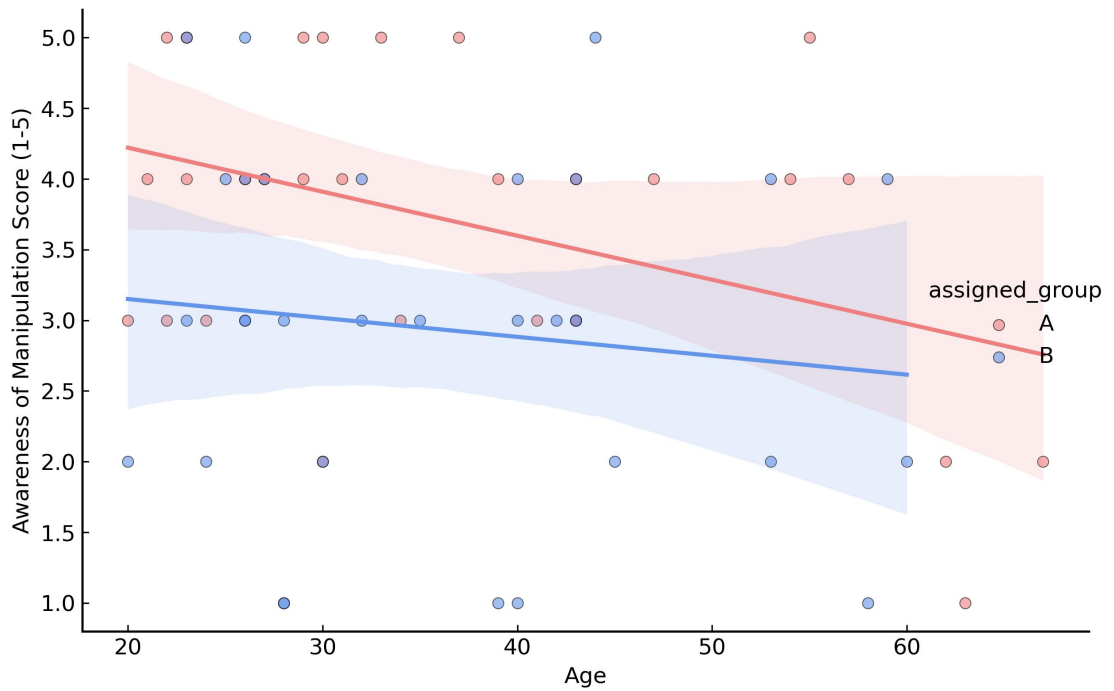


Figure 4.23 Correlation between Age and Awareness of manipulation score (q6).

Impact of Sex on Subjective perceptions

Independent-samples t-tests were conducted to compare the survey responses of male and female participants within each experimental group. Descriptive statistics for all subgroups are presented in Tables 4.27 and 4.28.

Table 4.27 Descriptive statistics for Perceived control (q3) and Trust (q4) by gender.

	Group A (Dark)		Group B (Ethical)	
Statistic (Score)	Male	Female	Male	Female
Perceived Control (q3)				
Mean (Std. Dev.)	3.90 (1.33)	4.11 (0.93)	4.23 (0.75)	4.11 (0.60)
Website Trust (q4)				
Mean (Std. Dev.)	3.20 (1.15)	3.11 (1.17)	3.50 (1.06)	3.89 (0.78)

Table 4.28 Descriptive statistics for Pressure (q5) and Awareness (q6) by gender.

	Group A (Dark)		Group B (Ethical)	
Statistic (Score)	Male	Female	Male	Female
Feeling of Pressure (q5)				
Mean (Std. Dev.)	2.65 (1.42)	3.11 (1.45)	2.00 (1.07)	3.33 (1.32)
Awareness (q6)				
Mean (Std. Dev.)	3.65 (1.14)	3.78 (0.97)	2.86 (1.25)	3.11 (1.27)

No statistically significant differences were found between genders in Group A. In Group B, a statistically significant difference was found for 'Feeling of Pressure' (q5), where females (M=3.33) reported significantly higher scores than males (M=2.00), $t(29) = -2.69$, $p = 0.019$. No other differences were significant. The results are visualized in Figures 4.24 and 4.25.

Chapter 4. Results

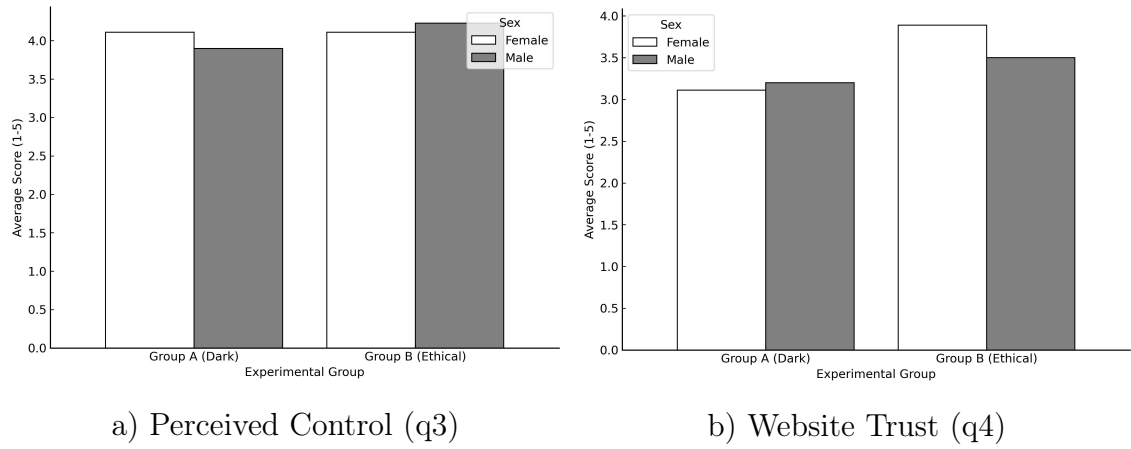


Figure 4.24 Comparison of Mean scores for Perceived control and Trust by gender.

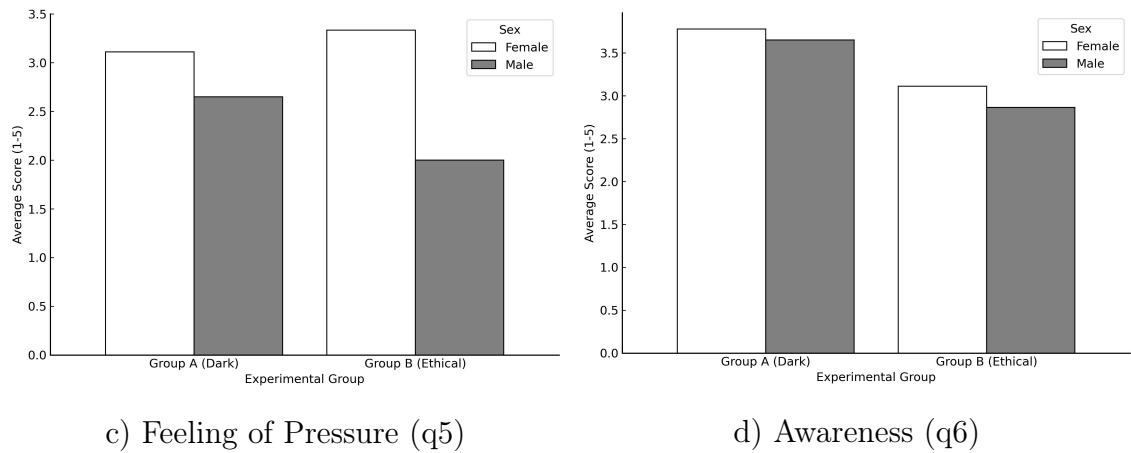


Figure 4.25 Comparison of Mean scores for Feeling of pressure and Awareness by gender.

4.3.3 Shame modal effectiveness (Group A)

The effectiveness of different shame modals was analyzed for users in Group A. Descriptive statistics detailing the outcomes of these encounters are presented in Table 4.29. A visual breakdown of these outcomes is shown in Figure 4.26.

Table 4.29 Descriptive statistics for Shame modal user outcomes (Group A).

Shame Modal Type	Convinced	Resisted	Total Encounters
Guest Checkout	1	17	18
Warranty	1	5	6
Post Checkout Offer	1	28	29
Save Card	3	6	9
Remove Item	0	1	1

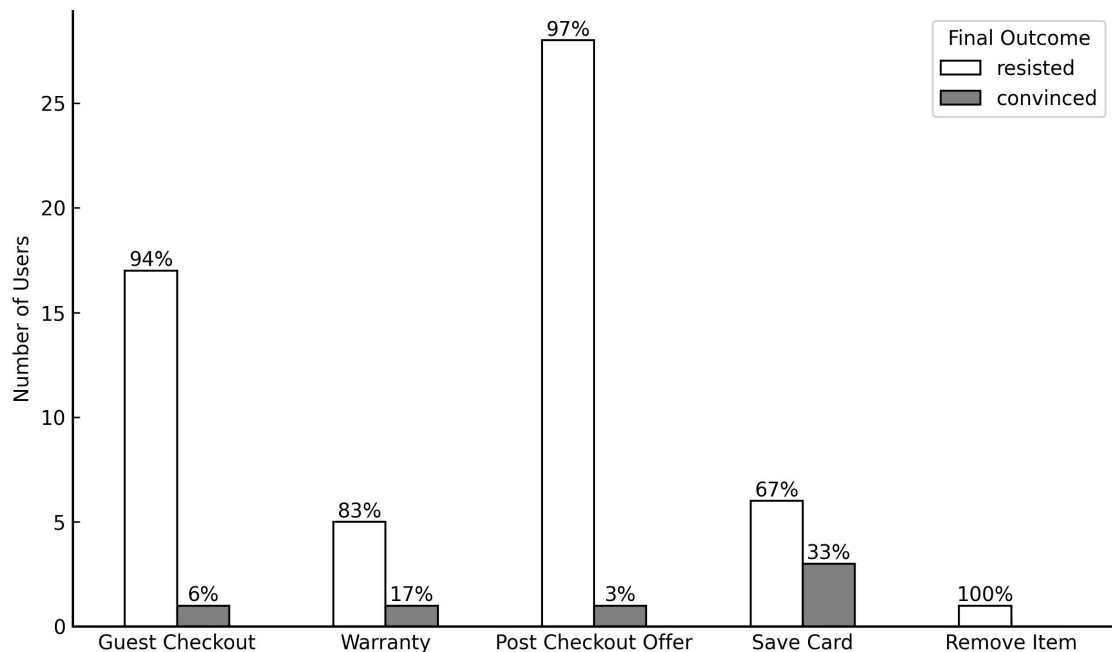


Figure 4.26 Effectiveness of Shame modals in Group A.

Chapter 5

Discussion

The experiment’s findings offer a complex and multidimensional perspective on the effects of dark patterns in a mock e-commerce setting. Some of the patterns that were put into practice produced unexpected and, in some cases, counterintuitive consequences, while others behaved just as predicted. This chapter analyzes these results by relating them to the initial research questions, discussing the wider ramifications for business and design, acknowledging the study’s inherent limits, and making recommendations for other research directions.

5.1 Interpretation of key findings

The data reveals a nuanced relationship between dark patterns and user behavior. The manipulative techniques largely succeeded in guiding users toward specific, pre-defined actions, such as accepting warranties. However, they failed to achieve the broader business goal of increasing overall purchase value, and they introduced significant friction that slowed the user journey. This study’s most notable contribution is highlighting the gap between a user’s conscious awareness of manipulation and their subsequent behavior and perception of control.

The ineffectiveness on overall spending

A central and unexpected finding was that the manipulative interface (Group A) failed to generate a statistically significant increase in the **Final order value** or **Subtotal** compared to the ethical interface. This challenges the common business assumption that implementing a suite of dark patterns will directly lead to higher overall revenue. This result may be partly explained by the simulated nature of the experiment, where the absence of real financial risk can alter purchasing behavior. However, it also suggests that the cognitive friction introduced by the dark patterns, such as the randomized product grid or the confirmshame modals, may have had a counteractive effect, discouraging the deeper browsing that could have lead to larger purchases.

The nuanced efficacy of behavioural nudges

The study revealed a clear distinction between the effectiveness of different types of manipulative nudges. Passive, low-friction dark patterns that exploited defaults were highly effective. By pre-selecting options for warranties, checkout protection, and saving card details, the interface successfully leveraged the status quo bias, a cognitive shortcut where people prefer to keep things as they are [14]. This resulted in significantly higher conversion rates for these items in Group A, indicating that users are susceptible to manipulation that requires no active decision-making.

In stark contrast, the active, high-friction dark pattern of **confirmshaming** was largely ineffective. As shown in Table 4.29, the shame modals were overwhelmingly resisted by users. For instance, the modal designed to prevent users from opting out of guest checkout was resisted by 17 out of 18 users who encountered it (94%). This suggests that while users may tolerate subtle manipulation, they actively push back against overt attempts to guilt or pressure them. This behavior aligns with the theory of **psychological reactance**, which states that when individuals feel their freedom of choice is being threatened, they will often assert their autonomy by doing the opposite of what they are being pressured to do [15]. The "Save Card" modal was a minor exception, with a slightly higher success rate (33%), perhaps because the perceived benefit of convenience was greater than the perceived risk of saving

fake card details.

The Awareness paradox: Seeing the trap but still stepping in it

Possibly the most notable finding was the disconnect between participants awareness of manipulation and their behavior. Participants in Group A were consciously aware that the site was trying to influence them, as shown by their significantly higher scores for **Awareness of manipulative elements** (Q6). Yet, this awareness did not translate into a statistically significant feeling of lost control (Q3) or trust (Q4), nor did it correlate with any change in their final spending.

This suggests that awareness alone is not a sufficient defense against many dark patterns. This phenomenon can be understood through dual-process theories of cognition, which separate intuitive "System 1" thinking from deliberate "System 2" thinking [16]. The most effective dark patterns in this study were the pre-selections and defaults, that targeted the automatic, low-effort System 1, which operates even when our analytical System 2 recognizes the manipulation. Users saw the trap, but the cognitively easier choice was to accept the default.

The vulnerability of older users

The demographic analysis provides an important, albeit exploratory, finding that aligns with the established body of research on older adults increased vulnerability to online deception and manipulation. The data revealed that older participants in the dark pattern group reported significantly higher levels of trust and control. Suggesting that this demographic may misinterpret manipulative cues, as the pre-selected "recommended" option, to be helpful guidance rather than a deceptive tactic. This vulnerability has real-world consequences, as confirmed by reports from consumer protection agencies that consistently show older adults suffer disproportionately high financial losses from online fraud [17]. This study's finding, therefore, provides direct experimental evidence within an HCI context for this well-documented vulnerability, indicating that dark patterns may be disproportionately effective and harmful for this group.

5.2 Implications and recommendations

The results of this study have practical implications for both designers and businesses.

- **For businesses:** The notion that a more aggressive, manipulative interface is a more profitable one is challenged. While specific add-ons can be boosted through pre-selection, the associated friction may suppress overall spending. Furthermore, the strong user resistance to overt confirmshaming suggests that such tactics are not only unethical but may also be ineffective, serving primarily to damage user trust.
- **For designers:** The "awareness paradox" places a heavy ethical responsibility on the designer. If users cannot be relied upon to protect themselves even when they notice manipulation, then designers must act as the primary defenders of user interests. This reinforces the need for a professional ethic that prioritizes transparency and user autonomy over exploiting cognitive biases.

5.3 Study limitations

While this study yielded valuable insights, it is important to acknowledge its limitations to better contextualize the findings.

- **Ecological validity:** The biggest limitation is that the experiment was a simulation. Participants were not spending their own money, which drastically lowers the stakes of the purchasing decision and is the most likely explanation for the lack of a significant difference in spending metrics.
- **Sample size and demographics:** While $N=60$ is adequate for detecting strong effects, it may not have been large enough to detect more subtle differences in subjective perceptions. Furthermore, the participants were recruited from Prolific and may not be fully representative of the general online shopping population.
- **Task design:** The instruction to "purchase at least one item" may have influenced participants to complete the task as quickly as possible rather than

engaging in genuine, exploratory browsing, as witnessed by the low rate of Products clicked in Figure 4.8 across both groups.

5.4 Future work

The outcomes and limitations of this study open up several compelling avenues for future research. The artificiality of spending fake money remains a key challenge in this type of HCI research. While a full-scale study with a real store is often infeasible, future work could explore this area with greater ecological validity without requiring a massive budget. Some potential directions include:

- **Qualitative "think-aloud" studies:** To better understand the "awareness paradox," a qualitative study could ask participants to speak their thoughts aloud as they interact with the dark and ethical interfaces. This would provide direct insight into their cognitive processes, helping to uncover why they notice manipulation but are still susceptible to it.
- **Longitudinal studies on trust and loyalty:** A study that invites participants back to the same experimental store a week or a month later could measure the long-term effects of their initial experience. This could test the hypothesis that while a dark pattern interface might achieve short-term goals, it erodes trust and reduces the likelihood of a user becoming a repeat customer.
- **Targeted demographic research:** The finding that older users may be more vulnerable warrants a dedicated study. Research focusing specifically on the mental models and digital literacy of different age groups could help explain why certain patterns are more effective on some demographics than others.
- **Cross-platform analysis (Mobile vs. Desktop):** The smaller screen size, touch-based interaction, and often more hurried context of mobile shopping may make users even more susceptible to dark patterns. An experiment that directly compares user behavior on mobile versus desktop versions of the same manipulative interface could produce results relevant to modern e-commerce design.

Chapter 6

Conclusion

This thesis set out to empirically measure the impact of dark patterns on user behavior and perception within an e-commerce context. Motivated by the widespread use of these manipulative techniques and the ongoing debate about their efficacy and ethical implications, this study compared two purpose-built e-commerce interfaces: one designed with a suite of common dark patterns and another based on transparent, ethical design principles. Through a between-subjects HCI experiment with 60 participants, objective interaction metrics and subjective survey responses were collected to provide a quantitative and qualitative answer to the central research questions.

The results of this experiment confirm that dark patterns are a powerful tool for influencing specific user actions. The dark pattern interface was significantly more successful at driving conversions on discrete, binary choices, such as account registration, the purchase of optional warranties, and the saving of payment details. However, this success did not translate into the primary business goal of increased overall spending, as there was no significant difference in the final order value between the two groups. Furthermore, the manipulative design came at a measurable cost to the user experience, resulting in a significantly longer and less efficient shopping journey for participants in the dark pattern group.

Chapter 6. Conclusion

Perhaps the most noteworthy contribution of this research is the identification of an "awareness paradox". Participants were consciously aware that they were being subjected to manipulative design, yet this awareness did not lead to a reported loss of control and trust, nor did it correlate with a change in their purchasing behavior. This suggests that the effectiveness of many dark patterns lies in their ability to exploit subconscious cognitive biases.

In conclusion, this study indicates that the business case for dark patterns may be less compelling than assumed. While they can successfully nudge users toward narrow, short-term goals, they do so by introducing friction and do not necessarily lead to greater overall revenue in a simulated environment. The findings place a heavy ethical responsibility on designers and businesses to prioritize user autonomy and transparency, as the research demonstrates that awareness alone is an insufficient defense against manipulative design. This work underscores the critical need for continued research and a greater industry focus on building user trust through ethical, user-centered design practices.

Bibliography

- [1] T. I. Hasan, C. I. Silalahi, R. Y. Rumagit, and G. D. Pratama, “Ui/ux design impact on e-commerce attracting users,” *Procedia Computer Science*, vol. 245, pp. 1075–1082, 2024.
- [2] R. Gunawan, G. Anthony, N. Vendly, and M. S. Anggreainy, “The effect of design user interface (ui) e-commerce on user experience (ux),” in *6th International Conference on New Media Studies (CONMEDIA)*, 2021.
- [3] B. J. Fogg, *Persuasive Technology: Using Computers to Change What We Think and Do*. San Francisco, CA, USA: Morgan Kaufmann Publishers, 2003.
- [4] A. Almutairi and M. Alshamari, “Assessment of persuasive design principles in e-commerce applications based on users’ reviews,” *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 9, no. 4, pp. 120–125, 2020.
- [5] H. Brignull. (2010, July) Dark patterns: Deception vs. honesty in ui design. Accessed: August 20, 2025. [Online]. Available: <https://www.darkpatterns.org/>
- [6] A. Mathur, G. Acar, M. Friedman, E. Lucherini, J. R. Mayer, M. Chetty, and A. Narayanan, “Dark patterns at scale: Findings from a crawl of 11k shopping websites,” *CoRR*, vol. abs/1907.07032, 2019. [Online]. Available: <http://arxiv.org/abs/1907.07032>
- [7] R. B. Cialdini, *Influence: The Psychology of Persuasion*, revised edition ed. Collins, 2007.
- [8] A. Tversky and D. Kahneman, “The framing of decisions and the psychology of choice,” *Science*, vol. 211, no. 4481, pp. 453–458, 1981.
- [9] J. Nielsen. (2012, January) Usability 101: Introduction to usability. Accessed: August 20, 2025. [Online]. Available: <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>

Bibliography

- [10] R. Likert, “A technique for the measurement of attitudes,” *Archives of Psychology*, vol. 140, pp. 1–55, 1932.
- [11] How Many Participants for Quantitative Usability Studies: A Summary of Sample-Size Recommendations. , September 2025. [Online]. Available: <https://www.nngroup.com/articles/summary-quant-sample-sizes>
- [12] J. W. Tukey, *Exploratory Data Analysis*. Reading, MA: Addison-Wesley, 1977.
- [13] H. B. Mann and D. R. Whitney, “On a test of whether one of two random variables is stochastically larger than the other,” *The Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 50–60, 1947.
- [14] R. Adkisson, “Nudge: Improving decisions about health, wealth and happiness, r.h. thaler, c.r. sunstein. yale university press, new haven (2008), 293 pp,” *The Social Science Journal*, vol. 45, p. 34–35, 12 2008.
- [15] J. W. Brehm, *A theory of psychological reactance*. Academic Press, 1966.
- [16] D. Kahneman, *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011.
- [17] Federal Trade Commission, “Protecting older consumers 2020-2021: A report of the federal trade commission,” Federal Trade Commission, Tech. Rep., October 2021. [Online]. Available: <https://www.ftc.gov/system/files/documents/reports/protecting-older-consumers-2020-2021-report-federal-trade-commission/protecting-older-consumers-report-508.pdf>

Abstract

While the existence and forms of dark patterns are well documented, their exact, measurable impact on user behavior and perception remains a critical question. This thesis addresses this gap through an HCI experiment in which participants were randomly assigned to one of two simulated e-commerce websites. Version A contained a suite of common dark patterns, including urgency, sneaking, misdirection, and social proof. In contrast, Version B was based on transparent and ethical design principles. To compare the two versions, objective interaction metrics (e.g. spending, session time, conversion rates) and subjective survey data on user perception were collected and statistically analyzed. The results show that although the dark patterns did not significantly increase the final order value, they were very effective in influencing discrete choices and led to significantly higher rates of account registration and purchase of optional add-ons. However, this came at the cost of a significantly longer total session time, suggesting increased friction in the interaction. An important finding was the "awareness paradox": participants in the dark pattern group were aware of the manipulative elements but did not report a significant loss of control, and this awareness did not correlate with their spending behavior. The results emphasize the importance of ethical design, as user awareness alone is not sufficient protection against cognitive exploitation.

Keywords — dark patterns, user experience, e-commerce, HCI experiment, ethical design

Sažetak

Iako su postojanje i forme tamnih obrazaca dobro dokumentirani, njihov točan, mjerljiv utjecaj na ponašanje i percepciju korisnika ostaje ključno pitanje. Ovaj rad adresira navedeno pitanje kroz HCI eksperiment u kojem su sudionici nasumično raspoređeni za testiranje jedne od dvije simulirane web stranice za e-trgovinu. Verzija A uključivala je skup uobičajenih tamnih obrazaca, uključujući hitnost, prikrivanje, obmanu i društveni dokaz. Nasuprot tome, verzija B zasnivala se na transparentnim i etičkim principima oblikovanja. Za usporedbu dvije verzije prikupljene su i statistički analizirane objektivne metrike interakcije (npr. potrošnja, trajanje sjednice, stope konverzije) i subjektivni podaci iz upitnika o percepciji korisnika. Rezultati pokazuju da iako tamni obrasci nisu značajno povećali ukupnu konačnu vrijednost narudžbe, bili su vrlo učinkoviti u manipuliranju diskretnim odlukama i doveli su do značajno viših stopa registracije računa i kupnje opcionalnih dodataka. Međutim, to je imalo cijenu u obliku značajno duljeg ukupnog trajanja sjednice, što sugerira povećano trenje u interakciji. Važno otkriće bilo je "paradoks svjesnosti": sudionici u skupini s tamnim obrascima bili su svjesni manipulativnih elemenata, ali nisu prijavili značajan gubitak kontrole, a ta svjesnost nije bila u korelaciji s njihovim potrošačkim ponašanjem. Rezultati naglašavaju važnost etičkog dizajna, jer sama korisnička svjesnost nije dovoljna zaštita od kognitivne eksploatacije.

Ključne riječi — tamni obrasci, korisničko iskustvo, e-trgovina, IČR eksperiment, etički dizajn