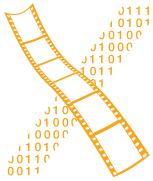


Chair for Computer Science 10 (Media Computing and Human-Computer Interaction)

RWTH AACHEN UNIVERSITY

Classify to Remember: Improving A Learning Game On Dark Patterns in User Interfaces

Bachelor's Thesis at the
Media Computing Group
Prof. Dr. Jan Borchers
Computer Science Department
RWTH Aachen University

by
Nikita Zaloga

Thesis advisor:
Prof. Dr. Jan Borchers

Second examiner:
Prof. Dr.-Ing. Ulrik Schroeder

Registration date: 26.07.2024
Submission date: 26.11.2024

Eidesstattliche Versicherung
Declaration of Academic Integrity

Name, Vorname/Last Name, First Name

Matrikelnummer (freiwillige Angabe)
Student ID Number (optional)

Ich versichere hiermit an Eides Statt, dass ich die vorliegende Arbeit/Bachelorarbeit/ Masterarbeit* mit dem Titel
I hereby declare under penalty of perjury that I have completed the present paper/bachelor's thesis/master's thesis* entitled

selbstständig und ohne unzulässige fremde Hilfe (insbes. akademisches Ghostwriting) erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt; dies umfasst insbesondere auch Software und Dienste zur Sprach-, Text- und Medienproduktion. Ich erkläre, dass für den Fall, dass die Arbeit in unterschiedlichen Formen eingereicht wird (z.B. elektronisch, gedruckt, geplottet, auf einem Datenträger) alle eingereichten Versionen vollständig übereinstimmen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.
independently and without unauthorized assistance from third parties (in particular academic ghostwriting**.** I have not used any other sources or aids than those indicated; this includes in particular software and services for language, text, and media production. In the event that the work is submitted in different formats (e.g. electronically, printed, plotted, on a data carrier), I declare that all the submitted versions are fully identical. I have not previously submitted this work, either in the same or a similar form to an examination body.

Ort, Datum/City, Date

Unterschrift/Signature

*Nichtzutreffendes bitte streichen/Please delete as appropriate

Belehrung:
Official Notification:

§ 156 StGB: Falsche Versicherung an Eides Statt
Wer vor einer zur Abnahme einer Versicherung an Eides Statt zuständigen Behörde eine solche Versicherung falsch abgibt oder unter Berufung auf eine solche Versicherung falsch aussagt, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.
§ 156 StGB (German Criminal Code): False Unsworn Declarations
Whosoever before a public authority competent to administer unsworn declarations (including Declarations of Academic Integrity) falsely submits such a declaration or falsely testifies while referring to such a declaration shall be liable to imprisonment for a term not exceeding three years or to a fine.
§ 161 StGB: Fahrlässiger Falscheid; fahrlässige falsche Versicherung an Eides Statt
(1) Wenn eine der in den §§ 154 bis 156 bezeichneten Handlungen aus Fahrlässigkeit begangen worden ist, so tritt Freiheitsstrafe bis zu einem Jahr oder Geldstrafe ein.
(2) Straflosigkeit tritt ein, wenn der Täter die falsche Angabe rechtzeitig berichtigt. Die Vorschriften des § 158 Abs. 2 und 3 gelten entsprechend.
§ 161 StGB (German Criminal Code): False Unsworn Declarations Due to Negligence
(1) If an individual commits one of the offenses listed in §§ 154 to 156 due to negligence, they are liable to imprisonment for a term not exceeding one year or to a fine.
(2) The offender shall be exempt from liability if they correct their false testimony in time. The provisions of § 158 (2) and (3) shall apply accordingly.

Die vorstehende Belehrung habe ich zur Kenntnis genommen:
I have read and understood the above official notification:

Ort, Datum/City, Date

Unterschrift/Signature

Contents

List of Figures and Tables

Abstract

Dark patterns are elements or tricks in user interfaces designed to manipulate the user into doing something, that is against their best interest. They occur throughout the internet and can lead to harmful consequences. Besides regulatory and technical countermeasures against dark patterns, an educational approach has been increasingly researched. Specifically, a learning game to educate the user to better detect manipulative patterns and gain confidence dealing with them has been created. This game employed constructed websites, where players had to first find dark pattern elements, and then classify them into one of five broadly defined categories. This classifying task was underdeveloped and less liked by players. In our thesis, we improve on it, by expanding the game with an existing, comprehensive ontology of dark patterns. This ontology features 65 different categories structured in a hierarchy of 3 levels. To employ it in the game, we design new menu interfaces for the selection of dark pattern categories. We implement them and conduct a user study to compare their usability. We measure timings for classifications of the added categories and interfaces, and record participants' opinions on the interfaces. We find, that menu interfaces employing the hierarchical structures of the ontology are most favored for learning dark patterns classification, and further, that including the ontology into the classifying task improves on the understanding of the workings of specific dark patterns after playing it. We also describe further research opportunities in regard to the implementation of the ontology and our menu interfaces in the learning game.

Überblick

Dark Patterns sind Elemente oder Tricks in Benutzeroberflächen, die darauf abzielen, den Benutzer zu manipulieren, damit er etwas tut, was nicht in seinem Interesse liegt. Sie kommen überall im Internet vor und können zu schädlichen Folgen führen. Neben regulatorischen und technischen Gegenmaßnahmen gegen dunkle Muster wird zunehmend auch ein pädagogischer Ansatz erforscht. Insbesondere wurde ein Lernspiel entwickelt, um dem Benutzer beizubringen, manipulative Muster besser zu erkennen und selbstbewusster mit ihnen umzugehen. Bei diesem Spiel wurden konstruierte Websites verwendet, auf denen die Spieler zunächst Dark Patterns finden und diese dann in eine von fünf allgemeiner definierten Kategorien einordnen mussten. Diese Klassifizierungsaufgabe war unterentwickelt und bei den Spielern weniger beliebt. In unserer Arbeit verbessern wir sie, indem wir das Spiel mit einer bestehenden, umfassenden Ontologie der Dark Patterns erweitern. Diese Ontologie umfasst 65 verschiedene Kategorien, die in einer Hierarchie von 3 Ebenen strukturiert sind. Um sie im Spiel zu verwenden, entwerfen wir neue Menüschnittstellen für die Auswahl der Dark Pattern Kategorien. Wir implementieren sie und führen eine Benutzerstudie durch, um ihre Benutzerfreundlichkeit zu vergleichen. Wir messen die Zeiten für die Klassifizierung der hinzugefügten Kategorien und Schnittstellen, und erfassen die Meinungen der Teilnehmer zu den Schnittstellen. Wir stellen fest, dass Menüschnittstellen, die die hierarchischen Strukturen der Ontologie verwenden, für das Erlernen der Klassifizierung von Dark Patterns am besten geeignet sind, und dass die Einbeziehung der Ontologie in die Klassifizierungsaufgabe das Verständnis für die Funktionsweise bestimmter Dark Patterns nach dem Spielen der Aufgabe verbessert. Wir beschreiben außerdem weitere Forschungsmöglichkeiten im Hinblick auf die Implementierung der Ontologie und unserer Menüschnittstellen in das Lernspiel.

Acknowledgments

First of all, I want to thank Prof. Dr Jan Borchers and Prof. Dr.-Ing. Ulrik Schroeder for examining this thesis.

Secondly, I want to sincerely thank my thesis advisers René Schäfer and Kevin Fiedler, for their insightful guidance and invaluable feedback during my work on this thesis. Your continuous support enabled me to stay motivated for the whole duration of the work.

I also want to thank all my friends and family who majorly supported me for the last couple of months. I want to especially thank my mother for keeping me encouraged, and my friends in the board of the Filmstudio at the RWTH, who took over many of my responsibilities in this time.

Finally, I would like to thank everyone who participated in my user study. Your detailed and extensive answers genuinely made the findings of this thesis more valuable.

Conventions

Throughout this thesis we use the following conventions:

The thesis is written in American English. The first person is written in plural form. Unidentified third persons are described with the pronouns they/their.

DEFINITION TEXT BOXES:
Definitions are set off in orange boxes.

Definition:
Definition Text Boxes

Where appropriate, paragraphs are summarized by one or two sentences that are positioned at the margin of the page.

This is a summary of a paragraph.

Chapter 1

Introduction

The internet has become ubiquitous in peoples' lives. In particular, mobile applications and websites are used extensively for almost any desired service. Communication, networking, information searching, entertainment in various forms, shopping, education, and more are accomplished or consumed over interactive, digital screens. Although most of the websites or apps offering these services are "free" to use, they are created with the intent to make money from their users [Maier and Harr, 2020].

Business goals make the exploitation of human psychology valuable.

As business goals are often valued over the user's best interests, many website designers deliberately exploit human psychological principles for monetary gains [Maier and Harr, 2020]. By using malicious design patterns, people can be nudged and manipulated into accepting terms, agreeing to hidden fees, or giving away their private data [Luguri and Strahilevitz, 2021]. The commonly used term for these patterns is "dark patterns" (or "deceptive patterns"), coined by Harry Brignull in 2010 [Brignull et al., 2023].

Definition:
Dark Patterns

DARK PATTERNS:
Dark patterns are elements or tricks in user interfaces, to make the user of a website do something they did not intend to and which may be against there best interest, based on their emotional or conditioned reactions. [Brignull et al., 2023; Maier and Harr, 2020]

Dark patterns occur not only on the most popular websites, but throughout the internet.

Most people probably get in contact with dark patterns regularly over the most visited websites like Google Search, YouTube, Instagram, X, and Amazon, which combine over 100 billion monthly visits worldwide [Similarweb, 2024]. These websites occur frequently in a list by Brignull et al. [2023], titled their "Hall of Shame" (https://www.deceptive.design/[1]), where incidents involving dark patterns are pointed out to the public. Overall, this list shows their commonness with over 500 examples throughout the internet.

Dark patterns can lead to harmful consequences, requiring for countermeasures to be explored.

As mentioned, dark patterns can be used to manipulate users into doing something they do not intend to. Besides causing mild annoyances, e.g., by making the user spend unnecessary much time on a website [Bhoot et al., 2020], consequences can be seriously harmful, like the user giving away more personal data than they wanted [Gunawan et al., 2022] or paying fees they didn't anticipate, leading to financial harm [Brignull et al., 2023].
Because of dangers like these, different countermeasures for dark patterns have been explored by lawmakers and researchers [Bongard-Blanchy et al., 2021]. Regulatory countermeasures have been instituted, including laws by the European Union or the Federal Trade Commission (FTC), for example when a company agreed to pay $18 million to settle charges[2] by the FTC for employing dark patterns. Also, technical countermeasures are being explored, with research revolving around detecting dark patterns and then highlighting or hiding them [Mathur et al., 2019; Schäfer et al., 2023, 2024]. In general, the efforts of human-computer interaction scholars to act against dark patterns continue relentlessly [Gray et al., 2024].

The educational approach is a user-focused type of countermeasure.

Besides the regulatory and technical countermeasures, there is another intervention measure that we can explore, focusing on the *education* of the user. Bongard-Blanchy et al. [2021] find that although users generally are aware of

[1] https://www.deceptive.design/hall-of-shame[Accessed: Nov.4,2024]

[2] https://www.ftc.gov/news-events/news/press-releases/2024/11/ftc-sends-more-17-million-consumers-harmed-brigits-deceptive-claims-junk-fees-confusing-cancellation[Accessed:Nov.5,2024]

the influence that manipulative designs can have on their behavior, they remain unsure of the actual workings of a dark pattern and its potential harm. By teaching the them, users may acquire the knowledge and confidence to detect and oppose dark patterns. Even if the detection is done by technical means, we assume that a knowledgeable user may be better equipped to withstand manipulation while using a website.
In their work on the educational approach, Bongard-Blanchy et al. [2021] furthermore suggest a "spot the dark pattern"-learning game that could teach the user about dark patterns.

LEARNING GAMES:
A learning game is a kind of serious game, that is not made with the exclusive intent of entertainment, and facilitates educational purposes. [Plass et al., 2015; Dörner et al., 2016]

Definition: *Learning Games*

There was only little research on the educational approach and no direct implementation of a general dark pattern learning game.

Learning games have proven to be a valuable approach to teaching about a specific topic [Griffiths, 2002]. However, there has been little research regarding this educational direction concerning dark patterns. In a recent contribution, Kronhardt and Gerken [2024] study the "effectiveness of persuasive games on behavioral changes in their players", applied to countermeasures against dark patterns, discussing principles on how to create a learning game for the topic of dark patterns. However, concrete implementations of such learning games are rare [Fiedler, 2024]. Besides a more abstract implementation by Kronhardt and Gerken [2024], a "fun to play" browser game[3], and another game focusing on preserving privacy data by Tjostheim et al. [2022], we found no dark pattern learning game intended for teaching the general detection of all types of dark patterns, like it was suggested by Bongard-Blanchy et al. [2021].

[3] https://cookieconsentspeed.run/[Accessed:Nov.5,2024]

Fiedler implements a concrete dark pattern learning game with strong results on players' recognition of and confidence dealing with dark patterns.

Fiedler [2024] attempts to "fill this gap" in their thesis and implements a fully functional learning game about recognizing and categorizing dark patterns on a constructed web page. They evaluate the effectiveness of their game in the improvement of the players' ability to detect dark patterns and the increase of their confidence dealing with them, by measuring the game's learning effects in a user study. Fiedler demonstrates, that, after playing the game, participants perform significantly better in detecting dark patterns and also were more confident in their decisions. Our thesis will focus on the further improvement of this learning game.

1.1 Motivation & Aim

There were only high-level pattern categories implemented in the classifying part of the game, which was generally less enjoyed than the finding part.

The game by Fiedler [2024] required the players to do 2 tasks: first, find a dark pattern on a constructed website, and then classify it into a selection of dark pattern categories. This classifying part only featured six high-level categories from a then preliminary ontology of dark patterns by Gray et al. [2023] (now updated in [Gray et al., 2024]). These high-level categories are rather abstract in their definitions, which may have led to participants "enjoying finding the dark patterns more than categorizing them". For example, a participant stated that *"Deciding the category is [...] annoying."* and another that *"It can be a bit frustrating to have a different view on what pattern applies to the situation."*

We want to improve on the classifying game task by including all categories and their hierarchical relations from Grays ontology.

The goal of our thesis is to improve on the aspect of classifying dark patterns in Fiedlers learning game, focusing on integrating all 65 categories from Grays ontology into the classifying task. We assume, that including more concrete, explanatory definitions for dark pattern types and their hierarchical relations may yield a better understanding of the workings and harms of specific dark patterns. Finally, we hope that this also leads to a beneficial learning effect for the overall recognition of and confidence dealing with dark patterns.

We will first explore, which menu interfaces are best suited for selecting dark pattern categories in the classifying task of the learning game by Fiedler [2024]. Then, we will present the results from a user study we conducted. We measured timings for selections made with these interfaces, and recorded quantitative and qualitative opinions on the different methods and interfaces employed. In the evaluation of the study, we aim to find results on the significance of including all dark pattern categories and their hierarchical relations in regard to achieving a better understanding of specific dark patterns.

We will evaluate a user study we conducted, to answer if a better understanding of dark patterns can be achieved by an improved classifying game part.

1.2 Outline

In Chapter 2, "Related Work", we will introduce the ontology of dark patterns by Gray et al. [2024] and compare it to other taxonomies. We further explore the domain of the educational approach to dark patterns and its goals, and specifically look into the structure of learning games. Also, we explore which designs for menu interfaces have proven to be viable in literature and popular applications.

In Chapter 3, "Exploring Menu Interfaces for the Classification of Dark Patterns", we present our design, implementation, and contribution of different menu interfaces for the classifying task in the dark pattern learning game by Fiedler [2024].

In Chapter 4, "User Study, Evaluation & Interpretation", we evaluate our implementations with a user study, including measured data and the participants' opinions, and then interpret our findings applied to the preceding research.

In Chapter 5, "Summary and Future Work", we conclude our research by summarizing our discoveries, noting their limitations, and suggesting future research opportunities and improvements for the game.

Chapter 2

Related Work

This chapter is divided into related work from three domains. Firstly, in Chapter 2.1, “Dark Patterns”, we look into the term of *dark patterns*, focus on taxonomies for them and explain how Gray et al. [2024] constructed their ontology for the categorization of dark patterns. Further, in Chapter 2.2, “The Educational Approach to Dark Patterns”, we explore the educational approach to dark patterns and its goals, look into the characteristics of learning games, and specifically examine the dark pattern learning game by Fiedler [2024]. Finally, in Chapter 2.3, “Viable Designs for Menu Interfaces”, we look into which designs for menu interfaces have proven to be viable in literature and popular applications, so that we can apply this knowledge to our own designs of menu interfaces for the selection of dark pattern categories.

2.1 Dark Patterns

Design patterns are recurring solutions to a problem.

The term *pattern* generally refers to "the repeated or regular way in which something happens or is done" (from Collins Dictionary[1]). In the field of human-computer interaction, the definition can be specified to "successful solutions to

[1] https://www.collinsdictionary.com/de/worterbuch/englisch/pattern[Accessed:Nov.8,2024]

recurring design problems", as Borchers [2000] state - however, they were referring to positive design patterns that describe good practices, whereas dark patterns describe established solutions to deceive and manipulate users [Bösch et al., 2016].

The term *dark pattern* was first used by Brignull et al. [2023], who started their website about dark patterns in 2010. Although specific definitions of the term vary in relation to particular domains, we can define them in general as follows:

Definition: *Dark Patterns*

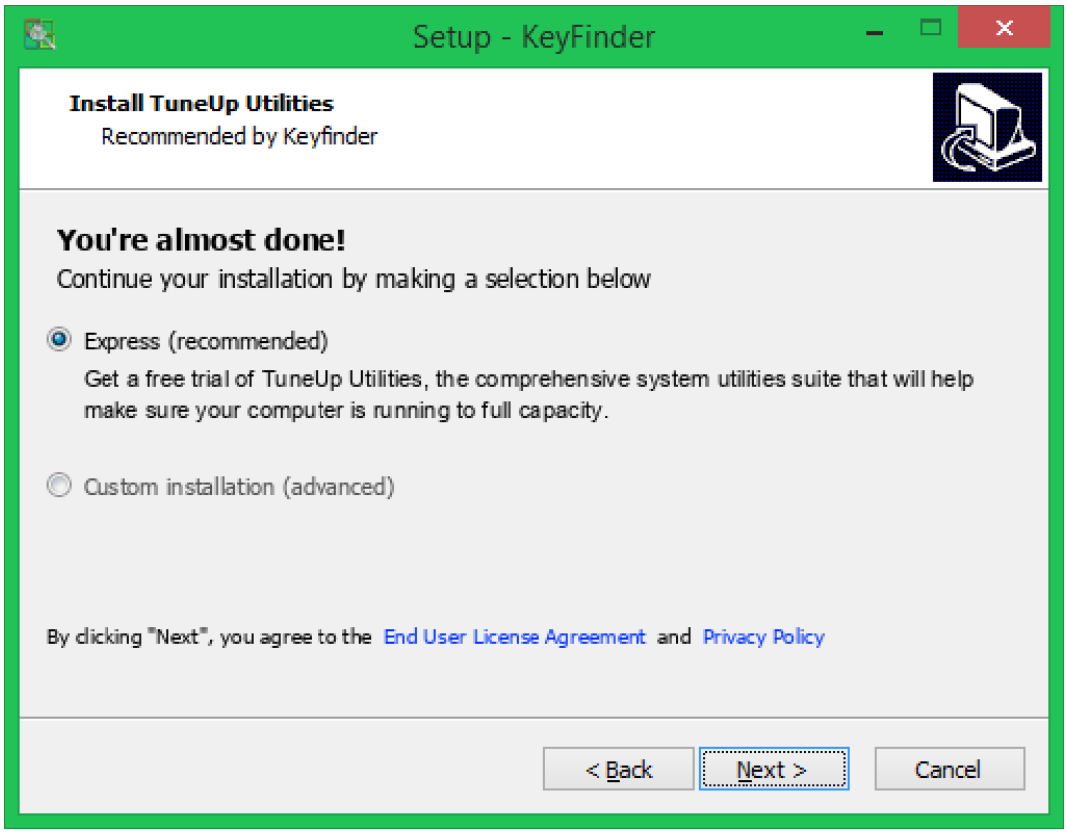
DARK PATTERNS:
Dark patterns are elements or tricks in user interfaces, to make the user of a website do something they did not intend to and which may be against there best interest, based on their emotional or conditioned reactions. [Brignull et al., 2023; Maier and Harr, 2020]

Figure 2.1 shows two examples of dark patterns.

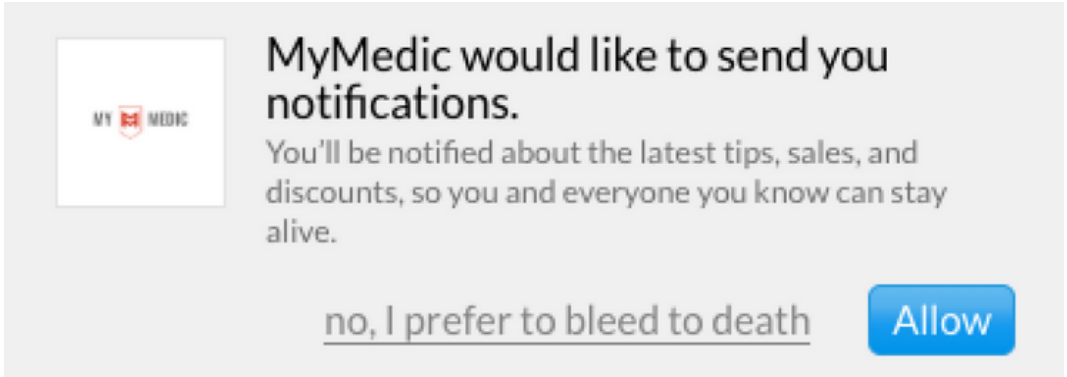
2.1.1 Taxonomies of Dark Patterns

Early taxonomies focus on high-level techniques or are limited in scope.

In April 2010, Conti and Sobiesk [2010] describe something that can be viewed as the predecessors to dark patterns: "malicious interfaces designs". Working on a foundation, they find 11 high-level techniques that deceptive interface utilize, like *confusion*, *distraction* and *obfuscation*. Further awareness is raised by Brignull et al. in July 2010 on their website [Brignull et al., 2023]. Originally listing only a few dark pattern categories, they expend these over the years to 16, including those from other taxonomies. They also rename dark patterns to "deceptive patterns".



(a) Screenshot from the Setup window of a program called "KeyFinder".



(b) Screenshot from a notifications pop-up on mimedic.com.

Figure 2.1: Two examples of dark patterns on real websites: (a) is an example of *False Hierarchy* (Interface Interference) [Gray et al., 2018], encouraging the user to think that one of the options is disabled by making it appear less important. (b) is an example of *Confirmshaming* (Social Engineering) [Brignull et al., 2023], inducing fear in the user that if they don't allow notifications, they will die.

Gray et al. [2018] expand on Brignull et al. [2023], introducing 5 high-level strategies.

In 2018, Gray et al. [2018] expand on Brignull et al. by proposing 5 high-level dark pattern strategies: *Nagging*, *Obstruction*, *Sneaking*, *Interface Interference*, and *Forced Action*. The existing dark pattern types by Brignull et al. are grouped into these. As Gray et al. consider strategic design decisions, similarities can be found between their high-level strategies and the high-level techniques by Conti and Sobiesk [2010].

Domain specific taxonomies are introduced.

Additionally, over the last decade, several domain specific taxonomies were introduced, including ones for: gaming [Zagal et al., 2013]; data privacy [Bösch et al., 2016]; shopping sites [Mathur et al., 2019], expanded upon in [Mathur et al., 2021]; and social networks [Mildner et al., 2023].

The Dark Pattern Ontology by Gray et al. [2024]

Gray et al. [2024] introduce an expansive ontology comprising existing taxonomies and regulatory reports.

Since 2018, Gray et al. made efforts to expand on their high-level strategies, working towards an expansive ontology of dark pattern types and their relations in [Gray et al., 2023]. This ontology is established in [Gray et al., 2024], and is the one we will work with in this thesis.
For the ontology, the authors build on Gray et al. [2018] and Brignull et al. [2023], aggregate the domain specific taxonomies mentioned before (expect for gaming dark patterns by Zagal et al.), include some additional dark pattern categories by Luguri and Strahilevitz [2021], and notably feature numerous regulatory reports from the EU, UK and USA. To account for the challenges of integrating these different sources, they worked in a team of established dark patterns researchers, including scholars with experience in human-computer interaction, design, web measurement, regulation, computer science and data protection law.

The ontology groups dark pattern types in a hierarchical structure of high-, meso- and low-level categories.

Gray et al. follow a refined methodology to create their ontology. First, they aggregate all types of dark patterns over all sources they included. They work separately for this step and find a total of 262 dark patterns. Then, they trace the origins of the patterns through direct citations and inferences to find commonalities. Following that, they decide together in the working group, which patterns are similar

and can be comprised into single categories, or don't fit the ontology and can be excluded.
Now, the categories the found patterns are grouped into are connected in a hierarchical structure with three levels. There are: *low-level*, *meso-level*, and *high-level* dark pattern categories defined.

High-level categories are more broadly defined.

The *high-level* categories are rather abstract, include general strategies for manipulative designs and are context independent, meaning they can be employed through numerous technologies (e.g. desktop, mobile...) and domains (e.g. shopping, gaming, social media...). Gray et al. [2024] essentially reuse their high-level strategies from Gray et al. [2018], however, *Nagging* moves to a lower level in the ontology and is replaced by *Social Engineering*, which mostly consists of added categories from Mathur et al. [2019].

Low-level categories are specific implementations.

The *low-level* categories are situational and dependent on context, including specific implementations or higher-level categories. For example, *Disguised Ads* is a low-level dark pattern category, defining all situations where interface elements are not clearly marked as advertisements (or otherwise biased sources) in a selection of interactive elements. It is an implementation of the high-level category *Sneaking*.

Meso-level categories bridge high- and low-level categories.

The *meso-level* categories aim to bridge high- and low-level categories together and describe an "angle of attack", which means a certain approach to implement a high-level dark pattern. They can be interpreted based on the specific context (technology and domain) of the dark pattern, resulting in a low-level category. The authors created them to either group low-level categories together, where the high-level category was too broad; or because a found dark pattern type was fitting the definition of a meso-level category better than a low- or high-level one. In that case, no additional low-level instantiations were defined for the new meso-level category.

In the same way as similar low-level categories are grouped together in a meso-level category, similar meso-level categories are grouped together in a high-level category, resulting in a hierarchical structure for the ontology. An example

of the high-level category *Interface Interference* and its subcategories can be seen in Figure 2.2. The finalized ontology by Gray et al. [2024] consists of 5 high-level, 25 meso-level and 35 low-level categories.

For each dark pattern category, Gray et al. [2024] created a definition matching a certain syntax.

Gray et al. [2024] created definitions for each dark pattern category. These match a respective syntax, depending on if it's for a high-, meso- or low-level category.

All high-level pattern definitions include an interplay of an *undesired action* and its *limitations* on the user:
HIGH-LEVEL DARK PATTERN is a strategy which UNDESIRED ACTION that [optionally, if known to users, would] DISTORT/SUBVERT/IMPEDE/OTHERWISE LIMIT USERS' AUTONOMY, DECISION-MAKING, OR FREE CHOICE.

All meso-level pattern definitions include a mismatch of the user's *expectations* and the actual *effect* resulting from an action:

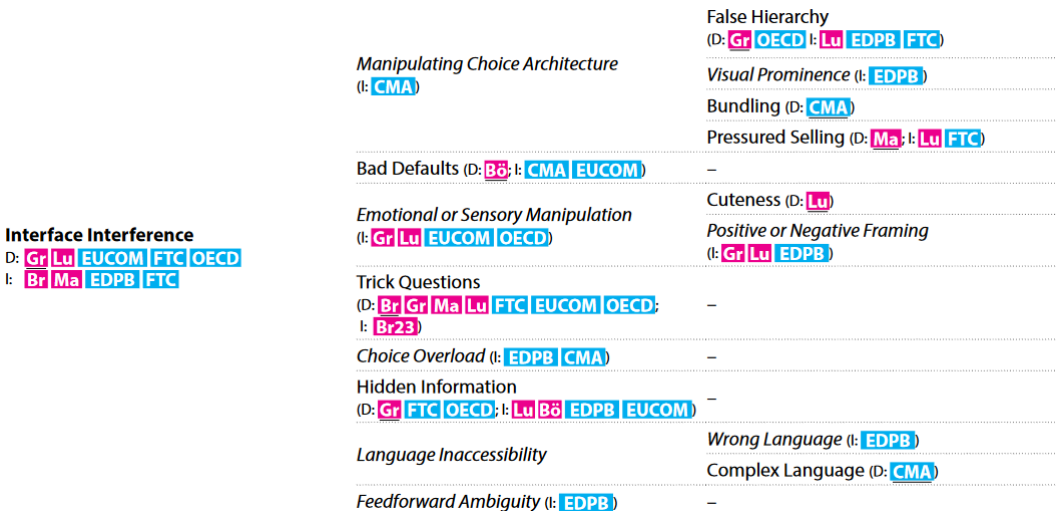


Figure 2.2: The high-level category *Interface Interference* and its subordinate meso-level categories (in column 2) and in turn, their respective low-level categories (in column 3) from the ontology by Gray et al. [2024]. The cyan colored abbreviations indicate that the pattern was found in a regulatory report, and the magenta ones that it was found in an academic source. "D" indicates a direct use of that pattern name in the original source(s) and "I" an inferred similarity between names used for two or more pattern types.

MESO-LEVEL DARK PATTERN subverts the user's expectation that EXPECTATION, instead producing or informing DIFFERENT EFFECT ON USER.

All low-level pattern definitions include an explanation, how their respective high- and meso-level patterns are *implemented* and what the *resulting effect* on the user and its *consequences* are.
LOW-LEVEL DARK PATTERN uses RELATED HIGH- AND MESO-LEVEL DARK PATTERN to ELEMENT OF UI ALTERED. As a result, INCORRECT USER EXPECTATION leads to UNDESIRED EFFECT ON USER.

The introduction of a sophisticated ontology for dark patterns raises new research questions.

The ontology by Gray et al. [2024] introduces a sophisticated collection of exemplary defined dark pattern categories and their relations. Applying this novel form of knowledge about dark pattern types to the educational approach to counteracting, particularly to learning games, may yield interesting results on the learning success of users. Also, the recognition of particular dark pattern types can be studied, as for example Bhoot et al. [2020] finds that some dark pattern types are recognized more often than others.

2.2 The Educational Approach to Dark Patterns

We focus on educating the user about dark patterns by utilizing learning games.

The educational approach to dark patterns aims to teach the user not only to be aware of the influence of manipulative designs, but also to acquire the knowledge and confidence to recognize and oppose dark patterns and their harms. As Bongard-Blanchy et al. [2021] point out, a knowledgeable user may be better equipped to withstand manipulation while using a website containing dark patterns. In this regard, our thesis focuses on educating the user, although there is research on educating the UX designers of apps and websites to not utilize dark patterns [Ahuja and Kumar, 2024; Caragay et al., 2024]. In particular, we look into *learning games* as a valuable approach to teaching.

However, the educational approach is not limited to learning games: For example, Lu et al. [2024] conduct a workshop to introduce participants to the topic of dark patterns.

Our educational goals are grounded on Blooms Revised Taxonomy (BRT), which arranges educational objectives in a two-dimensional array.

Regarding the goals of the educational approach, *Bloom's Revised Taxonomy* (BRT) by Krathwohl [2002] is widely adopted. In it, educational objectives are arranged in a two-dimensional system (see Figure 2.3 for a visualization). The first dimension is the *knowledge dimension*, including factual, conceptual, procedural, and metacognitive knowledge. The second dimension is the hierarchically structured *cognitive process dimension*, which includes (from simple to complex cognitive processes): remember, understand, apply, analyze, evaluate, and create. For our improvement of the learning game by Fiedler [2024], we are particularly interested in the *remembering* and *understanding* cognitive processes of the *conceptual* knowledge dimension, which are the educational objectives of *recognizing* and *classifying*.

Two anti-phishing learning games take BRT into account.

An example of a learning game relying on BRT is "Anti-Phishing Phil" by Sheng et al. [2007] - the authors describe in detail, how the game conveys both conceptual and procedural knowledge. A further example is another anti-phishing game by Röpke et al. [2022], who similarly explore, if higher levels on the cognitive process dimension in BRT can improve learning success. To achieve a higher level, they expand their classification game part, where URLs have to be classified into different categories. For teaching both conceptual and procedural knowledge types, games have proven to be effective [Gee, 2003].

2.2.1 Learning Games

Learning games can be defined as follows:

Definition: *Learning Games*

LEARNING GAMES:
A learning game is a kind of serious game, that is not made with the exclusive intent of entertainment, and facilitates educational purposes. [Plass et al., 2015; Dörner et al., 2016]

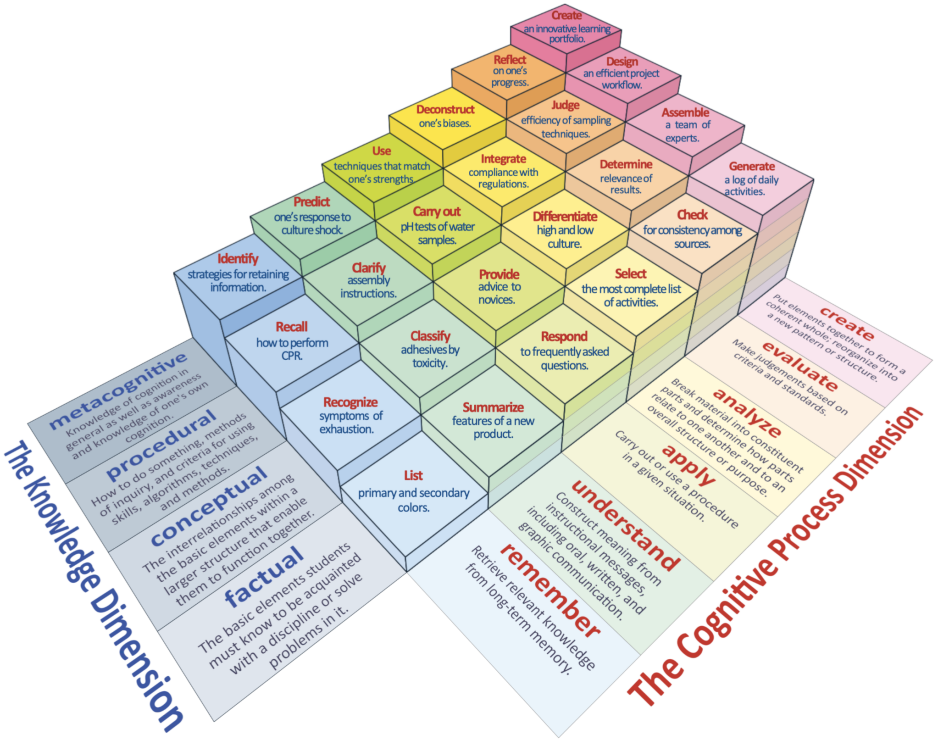


Figure 2.3: A Visualization of Bloom's Revised Taxonomy based on Krathwohl [2002]. The educational objectives are arranged in a two-dimensional system. Visualization by Rex Heer, Center for Excellence in Learning and Teaching, Iowa State University (licensed under a Creative Commons Attribution-ShareAlike 4.0 Internation License[1]).

[1] https://creativecommons.org/licenses/by-sa/4.0/[Accessed:Nov.14,2024]

Learning games are an engaging way to teach about specific topics.

Learning games are a valuable approach to teaching about a specific topic [Griffiths, 2002], extending knowledge that can be applied to everyday situations [Tang et al., 2009]. Dondlinger [2007] find, that they facilitate learning success for abstract thinking and deduction. They also motivate learners to engage with the topic longer [Yu, 2019; Tang et al., 2009] and can provide valuable feedback directly to the player [Tang et al., 2009].

An important aspect of the success of a learning game is its persuasive design.

In a systematic review of persuasive strategies in games, Ndulue and Orji [2022] show that persuasive games can lead to behavioral changes in the user. They analyze 130 games with a persuasive intent and find, that most of them are based on behavioral theories, like the model for persuasive design and behavior change by Fogg [2009]. For our case of learning games, we can understand our goal of "teaching" the user as a persuasive intent. Ndulue and Orji [2022] find, that positive effects on the behavior of players are generally persistent through persuasive games in various domains, including personal wellness, managing diseases, healthy eating and avoiding risky behaviors. Most notably, they observe that games with a persuasive intent are most effective when they employ *persuasive strategies*.

Utilizing persuasive strategies is important for a persuasive design.

These are defined in an exemplary way by Oinas-Kukkonen and Harjumaa [2009] in their framework for the design of persuasive systems, where they note 28 persuasive strategies in 4 categories. Examples of these strategies include *rehearsal*, implying a persuasive effect when the player has to repeat a certain behavior in the game; and *social comparison*, implying that players that can compare their performance to others may be more motivated to perform a learned target behavior.
Ndulue and Orji [2022] find that of these strategies *reward* is the most utilized, implying that a reward system in a persuasive game may more probably lead to change in behavior. Also, they find that the strategies in the category of *system credibility* are rarely used, proposing that presenting games as built upon scientific foundations may be helpful for persuasion intents. However, they argue that cognitive overload may occur when too many strategies are utilized at once, and that designers of persuasive games should focus on a few key strategies to implement.

Narratives in learning games facilitate learning success.

Naul and Liu [2019] find, that immersion, engagement, motivation, and learning success are increased in learning games, when a narrative is utilized. Narratives designed with certain characteristics like the inclusion of virtual agents and an intrinsically integrated story seem be the most effective in this regard.

Inoculation is the concept of exposing people to sufficiently weakened doses of something, so that a lasting immunity against it can be built (from Collins Dictionary[2]). The most prominent example of this are vaccines; however, we can also apply the concept to manipulative design elements. Saleh et al. [2021] actively apply inoculation to increase resistance against manipulative messages by extremist recruiters, and find it successful. They infer, that by exposing players of learning games to weakened doses of manipulative attacks (like in our case, dark patterns in a controlled, learning environment), an immunity to the manipulation of those can be achieved.

Inoculation theory can be applied to dark pattern learning games.

Until recently, there has been little research regarding learning games for dark patterns. Löschner and Pannasch [2024] measure the deceptive potential of certain dark patterns by letting participants play a decision-making game, but do not aim to teach them. Kronhardt and Gerken [2024] design a 3D narrative-driven learning game against deceptive patterns. In it, the player has to traverse different rooms that correspond to certain analogous dark pattern types, while a narrator confuses the user with the intent of making them stay longer than necessary in a room. For example, one room contains a hallway that is visually highlighted by light sources, while the shorter path through a hidden door remains dark, employing the metaphor of "visual prominence". The authors refer to inoculation [Saleh et al., 2021] and persuasive strategies [Oinas-Kukkonen and Harjumaa, 2009] as concepts they utilize for the design of their game.

Kronhardt and Gerken [2024] employ inoculation theory and persuasive strategies for their dark pattern learning game.

Kronhardt and Gerken [2024] conclude with principles for the general design of dark pattern learning games: in particular, they propose that game designers should decide on the persuasive goal of their game early in the design process. They note, that if the goal is to convey specific knowledge about dark patterns, they should ideally be depicted using real-world examples, with the aim of players categorizing examples when they are faced with them; and if the goal is to improve resistance against deceptive patterns (the authors' state this as the goal for their own learning

They note, that a game's persuasive goal should be defined early on.

[2] https://www.collinsdictionary.com/dictionary/english/inoculation[Accessed:Nov.16,2024]

game), they should be depicted more abstractly, with the aim of players recognizing that they are being manipulated or learning possible counter-measures.

Before Fiedler [2024], there was no direct implementation of a general dark pattern learning game.

Other, concrete implementations of dark pattern learning games are rare [Fiedler, 2024]. Besides the more abstract implementation by Kronhardt and Gerken [2024], a "fun to play" browser game[3], and another game focusing on preserving privacy data by Tjostheim et al. [2022], we found no dark pattern learning game intended for teaching the general detection of all types of dark patterns, like it was suggested by Bongard-Blanchy et al. [2021]. Fiedler [2024] attempt to "fill this gap" in their thesis, and implement a fully functional learning game about recognizing and categorizing dark patterns on a constructed web page. As our contributions are largely based on their work, it is described in detail in the following.

The Dark Pattern Learning Game by Fiedler [2024]

The learning game by Fiedler [2024] is based on a mechanic, where the player has to first find dark patterns on a website, and then classify them.

The dark pattern learning game by Fiedler [2024] is based on a game mechanic they call *Multi-Spot*. The player has to investigate a constructed website for dark patterns, select areas where they suspect a dark pattern, and finally assign it to a category from a selection of dark pattern categories. The game is developed with React[4], which employs modular elements, facilitating variable levels and fast content creation. The game also includes a narrative around the theme of the fictitious *"Dark Pattern Defense Force"*, which the player is recruited into. The design of the user interface supports the story, as for example, the levels are called "contracts".

The game consists of a tutorial and the main game levels.

Before the player can access the regular levels, they have to play a tutorial . Here, they are introduced to the topic of dark patterns and the original 6 high-level dark pattern types from Gray et al. [2023]. Each type is explained with one example. After a short introduction to the game mechanics, the player is quizzed on the information from the

[3] `https://cookieconsentspeed.run/[Accessed:Nov.5,2024]`
[4] `https://react.dev[Accessed:Nov.16,2024]`

tutorial, having to classify preselected elements into the 6 categories.
In the main game levels, the player has to select suspected dark pattern elements on the page with the *"Dark Pattern Magic Wand"*, which they can find in the *sidebar* on the right of the website (see Figure 2.4). After the selection, they have to assign the potential dark pattern element to one of the six high-level dark pattern types, for which corresponding buttons appear in the sidebar (see Figure 2.5). The player can also choose the option *"I'm not sure"*, or select another element on the page if they changed their minds. In the original version, 5 levels are included, themed around a smartphone online shop and a hotel booking site (the latter can be seen in Figure 2.4 and 2.5).

A feedback and a scoring system are included.

A feedback system is included: When the player categorizes an element, an overlay is presented directly. It shows whether the choice of the dark pattern type was correct, and a short explanation for the correct classification. Also, the element becomes colored in either red, yellow or green, depending on the classification being wrong, the player selecting the *"I'm not sure"* option, or the classification being correct, respectively. Falsely selected elements that are no dark patterns at all are also colored red. After finishing a level, the player can choose to see all potentially missed dark patterns also highlighted.
Players are scored for finding dark patterns and classifying them respectively. Score changes are showed in the feedback overlay, and a total score can be viewed at the top of the sidebar.

Fiedler [2024] conducts a user study to evaluate measurable learning effects and a potential increase of confidence in the player.

Fiedler [2024] conducts a user study to evaluate, if playing the game causes a measurable learning effect, and if the player gains confidence in detecting dark patterns by playing the game. The participants played the tutorial and all levels at their own pace. The learning effect was measured by a "pre-test/post-test": the participants took a test before playing the game, where they were shown images of websites and asked whether they were manipulative. After playing the game, the test was repeated with the same images and some additional ones. The game also utilizes questionnaires : Before the study, the participants had to fill out a first questionnaire requiring demographic data.

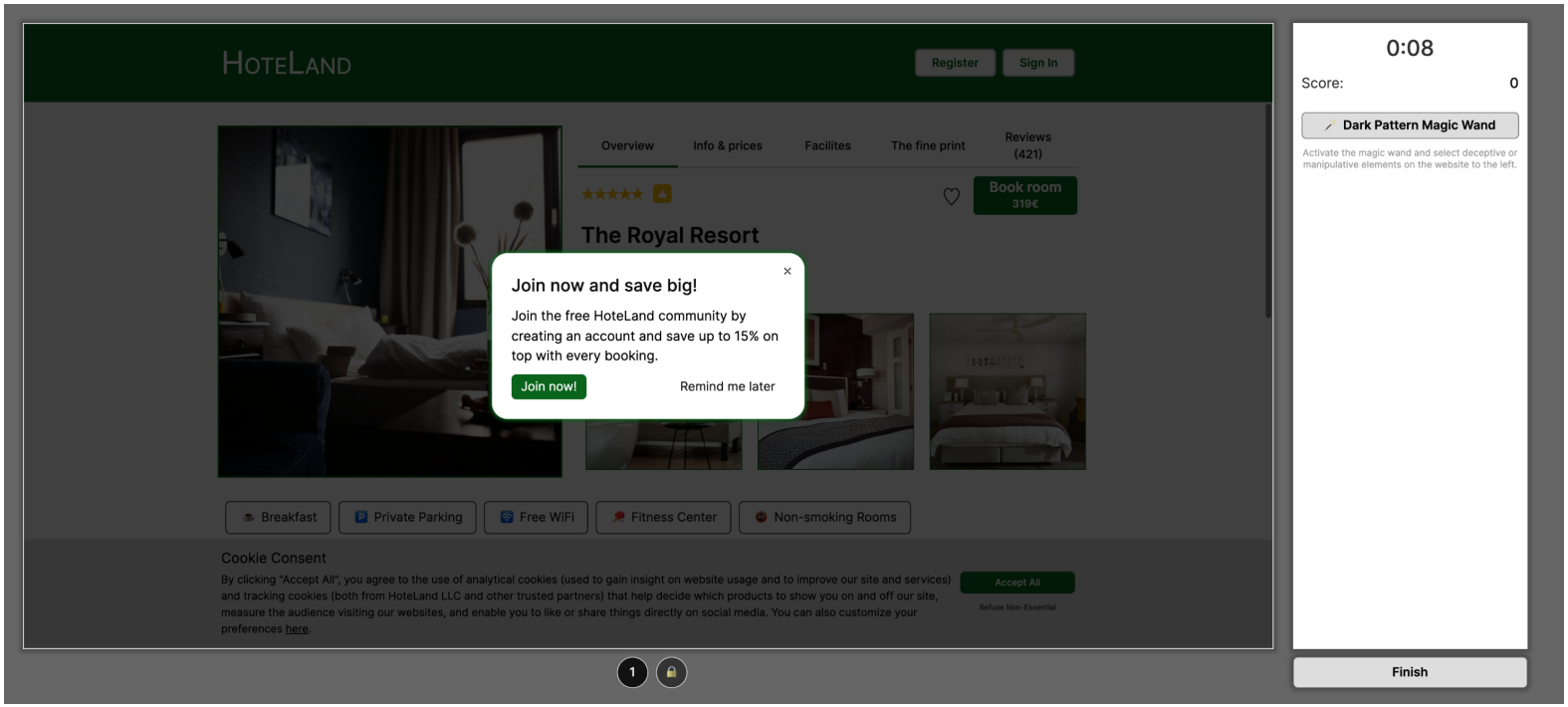


Figure 2.4: The view of a main game level by Fiedler [2024] *before* selecting an element. The sidebar on the right includes a timer, score display and a button for activating or deactivating the "Dark Pattern Magic Wand". The level can be finished with the "Finish" button at the bottom right corner.

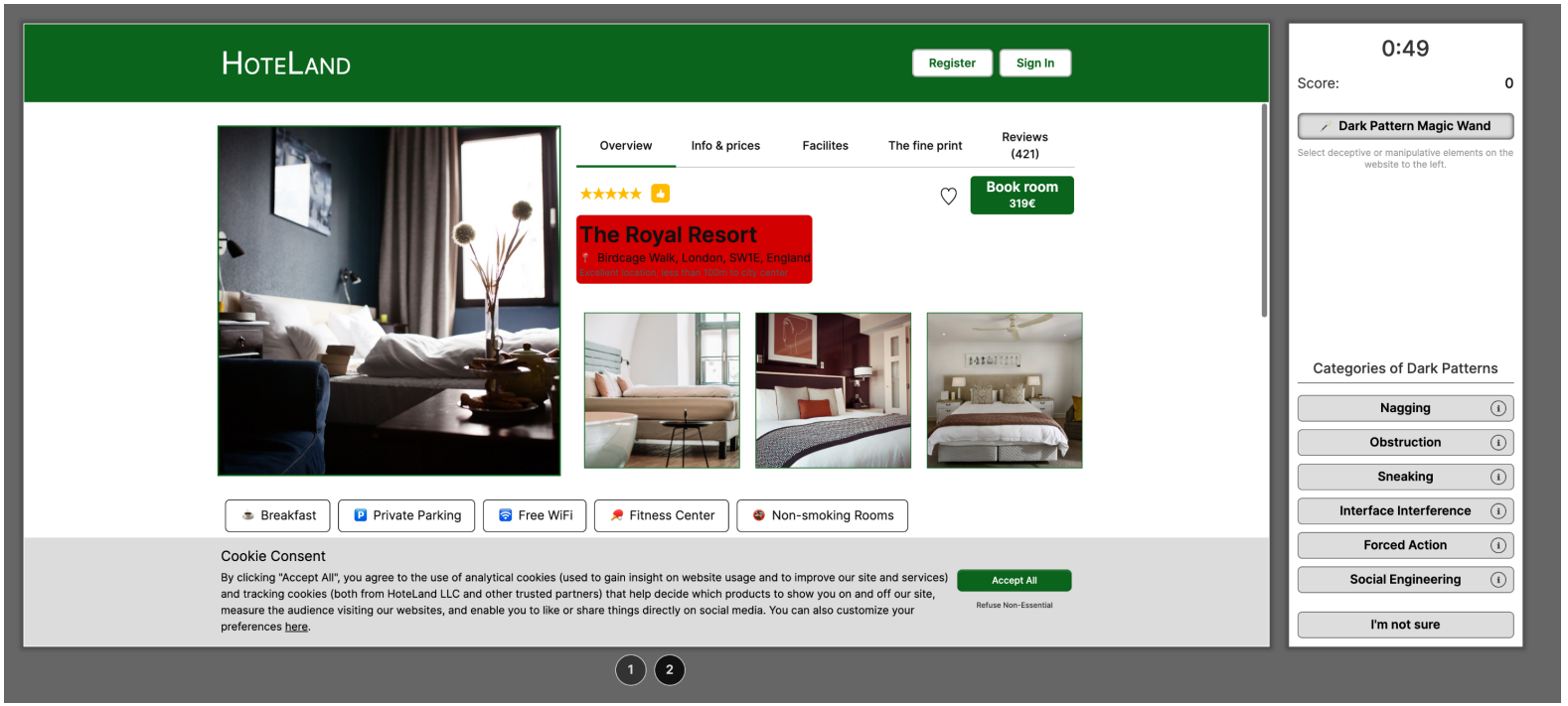


Figure 2.5: The view of a main game level by Fiedler [2024] *after* selecting an element. The element is highlighted in red. The sidebar on the right now additionally includes a selection of possible dark pattern category types, of which the player has to choose one.

A second questionnaire was employed after the study, to assess the participants opinions on the tutorial and game. The second questionnaire also included questions to measure the potential increase of confidence in recognizing dark patterns.

A learning effect and increase of confidence can be measured after playing the game. Also, the task of finding dark patterns is more liked than classifying them.

The author finds that there is a significant learning effect measurable for the performance of the participants in the repeated test after playing the game. Also, participants state, that playing the game increased both their knowledge about dark patterns as well as their confidence in detecting them, and to a lesser extent, also their confidence in handling the dark patterns. In a qualitative analysis of open questions asked in the questionnaires, Fiedler [2024] also finds that although the participants generally enjoyed finding the dark patterns, many disliked the game part of classifying them. For example, a participant stated that *"Deciding the category is [...] annoying."* and another that *"It can be a bit frustrating to have a different view on what pattern applies to the situation."* Notably, the participants were given only high-level categories as options in the classifying task.

2.3 Viable Designs for Menu Interfaces

> *"If Shakespeare could write 'all the world is a stage,' an interface designer could point to the computer screen and say 'all the interface is a menu.'"*
>
> *—Kent L. Norman*

There is extensive research on the design of menu interfaces for selection.

Since the invention of the first bit-mapped computer displays that made graphical user interfaces possible, extensive research for the design of graphical menu interfaces intended for the selection of items has been conducted [Callahan et al., 1988]. Menus for selection have been employed in operation systems, applications and games, both on mobile and computer devices [Norman, 2008].

There are some generally established characteristics that can be measured to evaluate the effectiveness of concrete implementations of menu interfaces, most notably: the time it takes to find an item; and the time it takes to navigate to an item to select it [Samp and Decker, 2010]. Considerations for the design of effective menus include: a meaningful order of the items selectable; how the items are visualized to be distinguishable, e.g. by using graphics like icons; and how the items are clustered or hierarchically organized. Regarding the latter, a hierarchical menu with more breadth than depth is generally better suited for quick selections [Norman, 2008]. In the following, we introduce established options for menu interfaces and note some of their advantages and disadvantages.

The effectiveness of concrete menu interfaces can be measured.

Linear, top-to-bottom listing menus have been predominant as menu interfaces at first [Callahan et al., 1988]. In their comparison of linear vs. pie menus, Callahan et al. [1988] find, that their main advantage lays in being flexible in the amount of displayed options, as the list of items can be extended variably by different methods. Firstly, the window size of the list can be limited and a scrolling mechanism added (e.g., a *scroll-bar*, see Figure 2.6a). Bederson [2000] explore this method in more detail, introducing their own implementation of a kind of scrolling mechanism they name the *"fisheye menu"* (see Figure 2.6b).

Linear menus have been widely employed and allow for a flexible amount of items.

Secondly, the list can be organized hierarchically. Zaphiris et al. [2001] explore, how a hierarchy of options can be accessed, comparing *expandable menus* with *sequential menus* (see Figure 2.6c for the latter), which differ in the amount of options from different hierarchical levels the user can see at once. One can argue that the *cascading menu* (see Figure 2.6d), which is widely employed in many applications, is a type of expandable menu [Cockburn and Gin, 2006]. Similarly, accordion menus[5] use the paradigm of an expandable menu.

Linear menus can also be structured and presented hierarchically.

[5] https://www.nngroup.com/articles/accordions-on-desktop/ [Accessed:Nov.18,2024]

Linear menus are well suited for query-based filtering.

Also, a list of options can be filtered by using query-based methods, which is especially useful for large amounts of hierarchically parallel options [Hochheiser and Shneiderman, 2000], e.g. for pages on a website. For example, a *searchbar* (see Figure 2.6e) can be used, in which the user can input text, which is then matched to the options displayed. However, query-based methods without additional navigational menus have its own disadvantages. For instance, they require more knowledge about the search space and increase cognitive load on the user, as [Budiu, 2014] point out in their article.

Linear menus are very common throughout operation systems and applications. Especially cascading menus can be found at the top of many windows, like in the menu bar of Visual Studio Code[6] or in Microsoft Office products[7].

Radial menus have been extensively compared to linear menus, demonstrating advantages in selection, but disadvantages in search times.

Another form of menu interfaces for selection are the *radial menus* (or *pie* menus). In contrast to linear menus, the options to select appear in a circle, typically around a spot the user has clicked or touched [Callahan et al., 1988; Samp and Decker, 2010]. Callahan et al. compare radial menus to linear ones and find, that radial menus reduce the time it takes to navigate to an item to select it. In another comparison by Samp and Decker [2010], the authors can replicate this finding. However, they also find that visual search is faster for linear menus (the time it takes to find an item), so in total, radial menus do not necessarily allow for a quicker selection than linear ones. Also, Callahan et al. [1988] find, that the radial menu can obscure the view of the element targeted by the user.

Radial menus can only comprise a limited amount of options, and have a limited amount of space for the options themselves.

A notable disadvantage of radial menus is, that they have a limited amount of options that can be placed around a circle. Although the number of options can be increased by enlarging the circle or shrinking the interactive options themselves, for too many options the interface becomes unusable [Samp and Decker, 2010]. In general, the options don't have as much space for text - which is inherently horizontally aligned - as options in linear menus, because the texts can overlap with each other in the circle. In regard

[6] `https://code.visualstudio.com/[Accessed:Nov.18,2024]`
[7] `https://www.office.com/[Accessed:Nov.18,2024]`

to this, Samp et al. note, that circular shapes are to be preferred for radial menus, as they can be placed conveniently in a circle. Furthermore, they state that icons are well suited for these circular shapes, instead of text boxes. However, Norman [2008] recommends that designs that use icons either include explanatory text with the icon, or display its function when the user is hovering over it with the mouse cursor, to make them more easily identifiable.

Like linear menus, radial menus can be structured and presented hierarchically.

In their comparison with linear menus, Samp and Decker [2010] further explore the implementation of hierarchically structured lists for radial menus. They propose *extending radial menus* (see Figure 2.6f) that work analogously to expanding linear menus, and more specifically, to cascading menus, which they employ for their comparison in a user study.

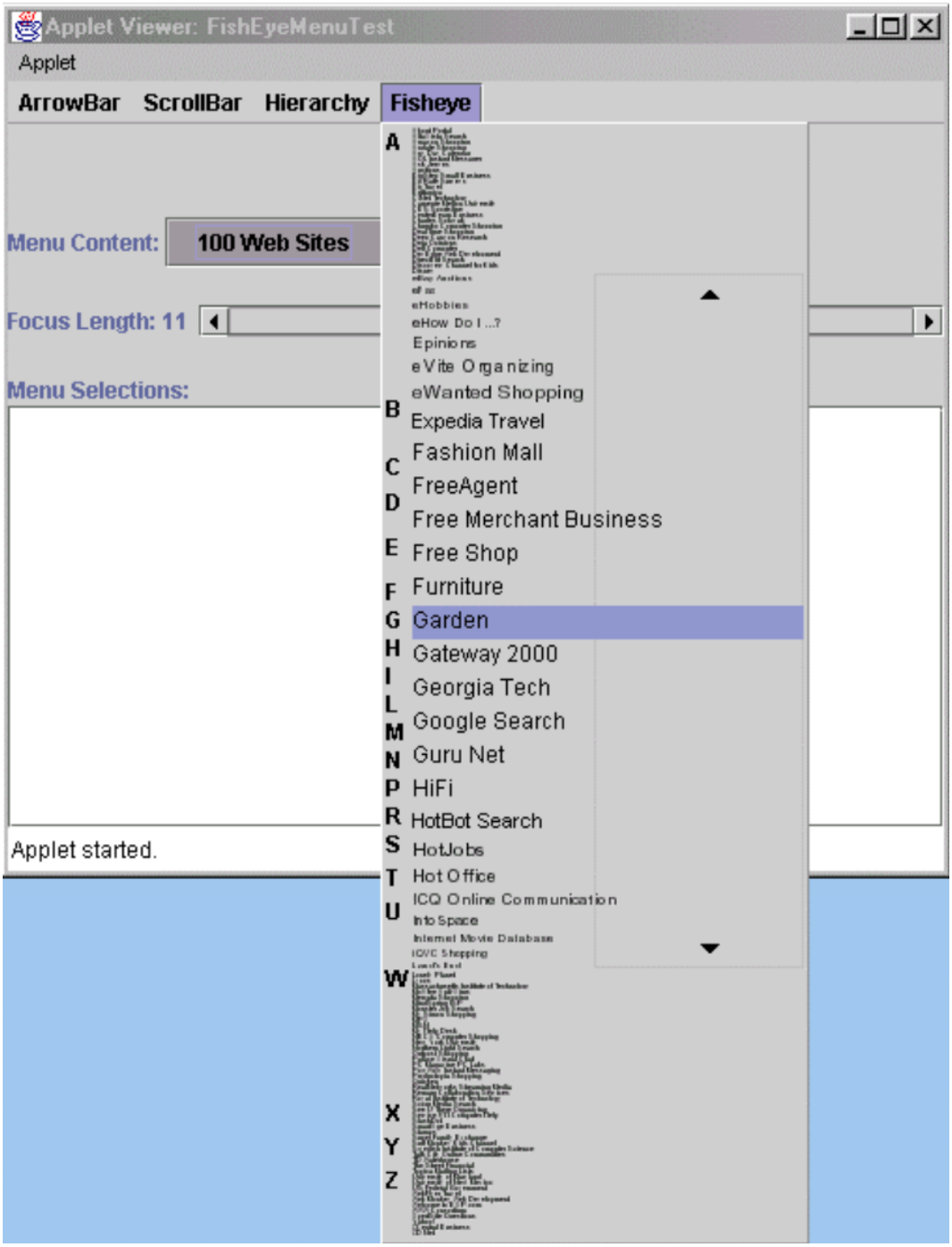
Radial menus can be found in a variety of popular applications. For example, ProCreate offers the QuickMenu[8], a customizable radial menu for frequently used functions.

[8] https://help.procreate.com/procreate/handbook/interface-gestures/quickmenu[Accessed:Nov.18,2024]

(a) A list menu extending outside the bounds of the window, making the options accessible with a *scroll-bar*. Screenshot from Microsoft Word[1].

[1] `https://www.office.com/[Accessed:Nov.18,2024]`

Figure 2.6: Different variants of existing menu interfaces

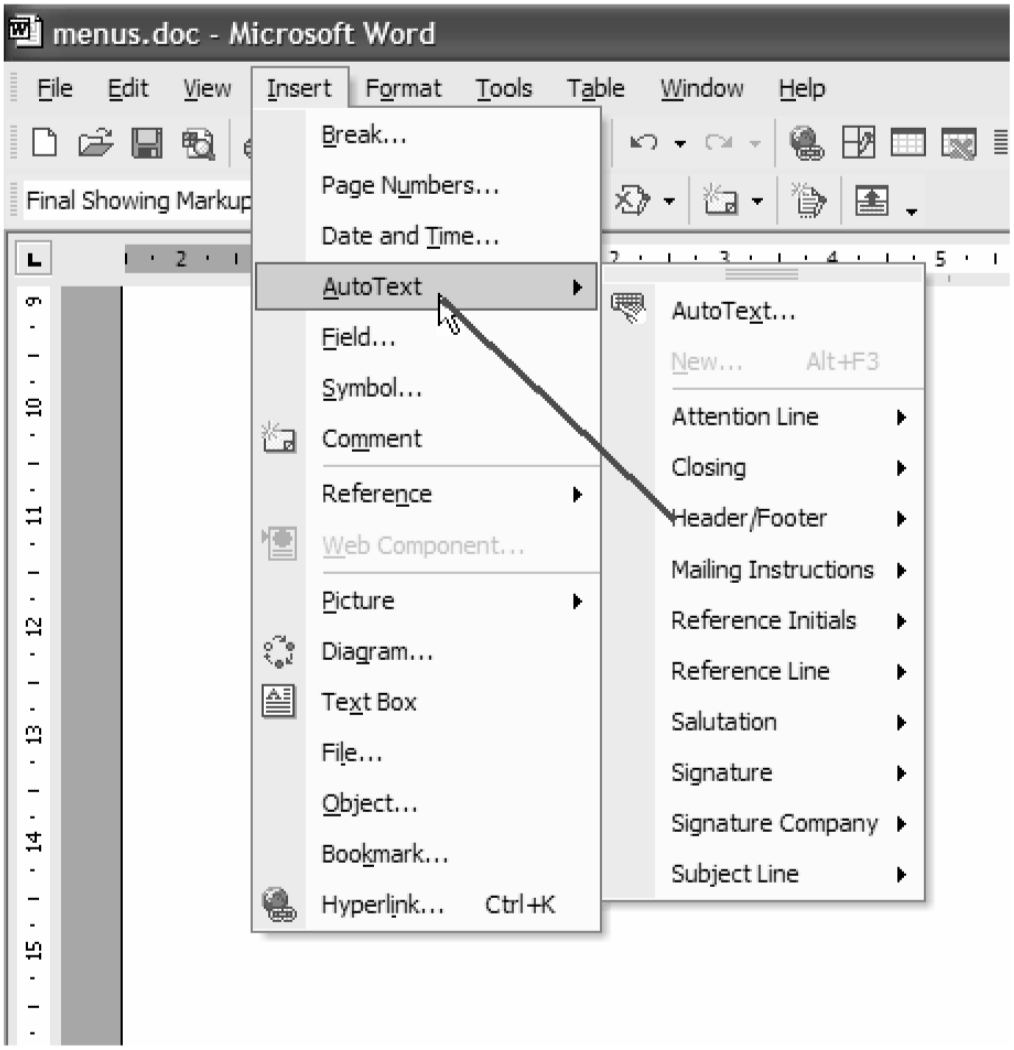


(b) The *fisheye menu* by Bederson [2000]. The interface works similar to a scrollbar, "zooming" in on the currently viewed options, instead of moving them into the view.

Figure 2.6: Different variants of existing menu interfaces (cont.)

(c) A depiction of a *sequential menu* by Hochheiser and Shneiderman [2000]. The user makes one choice after the other at each hierarchy level, replacing the previous selection options with the next. They can also choose to "return" to the previous menu or the first menu of the sequence. Note that usually in sequential menus, there is only one window and not multiple coexisting like in the figure.

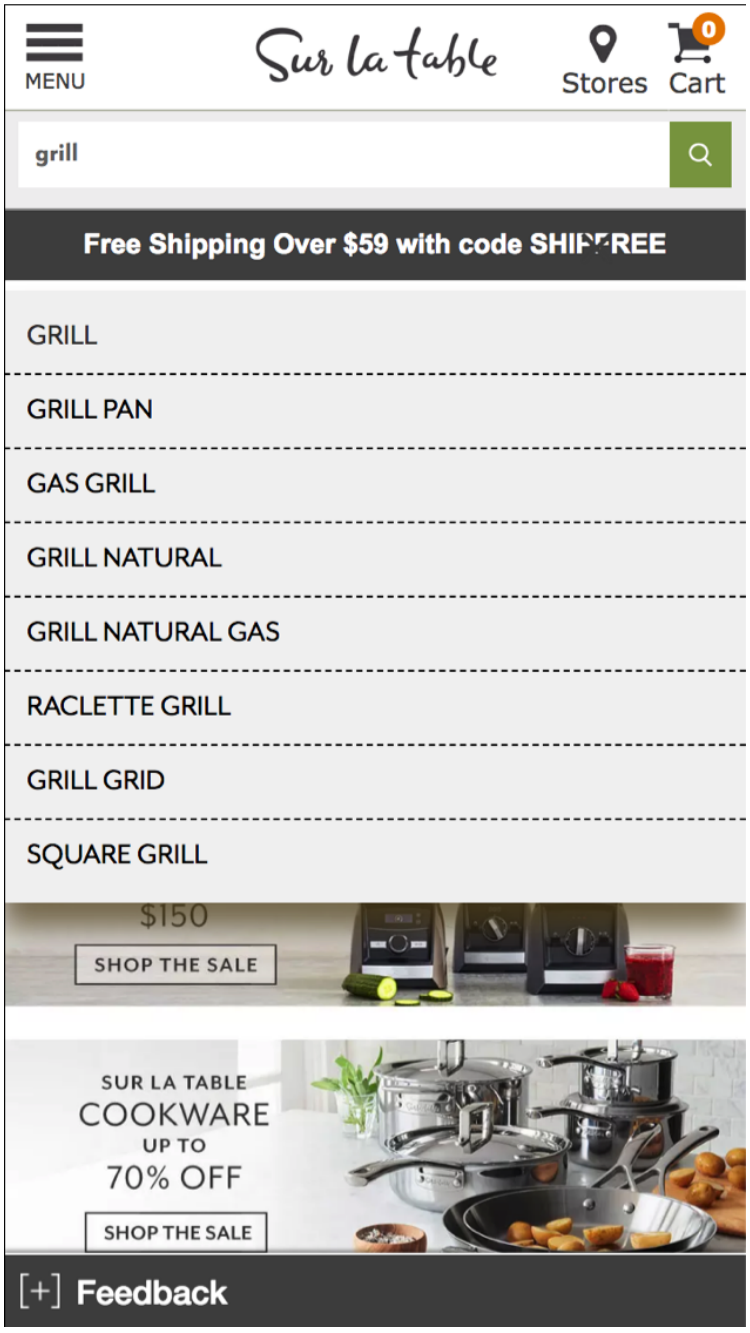
Figure 2.6: Different variants of existing menu interfaces (cont.)



(d) An example for a *cascading menu* by Cockburn and Gin [2006] from an older version of Microsoft Word[1]. Hovering over an option with an arrow symbol reveals further options at a lower level in the hierarchy.

[1] `https://www.office.com/[Accessed:Nov.18,2024]`

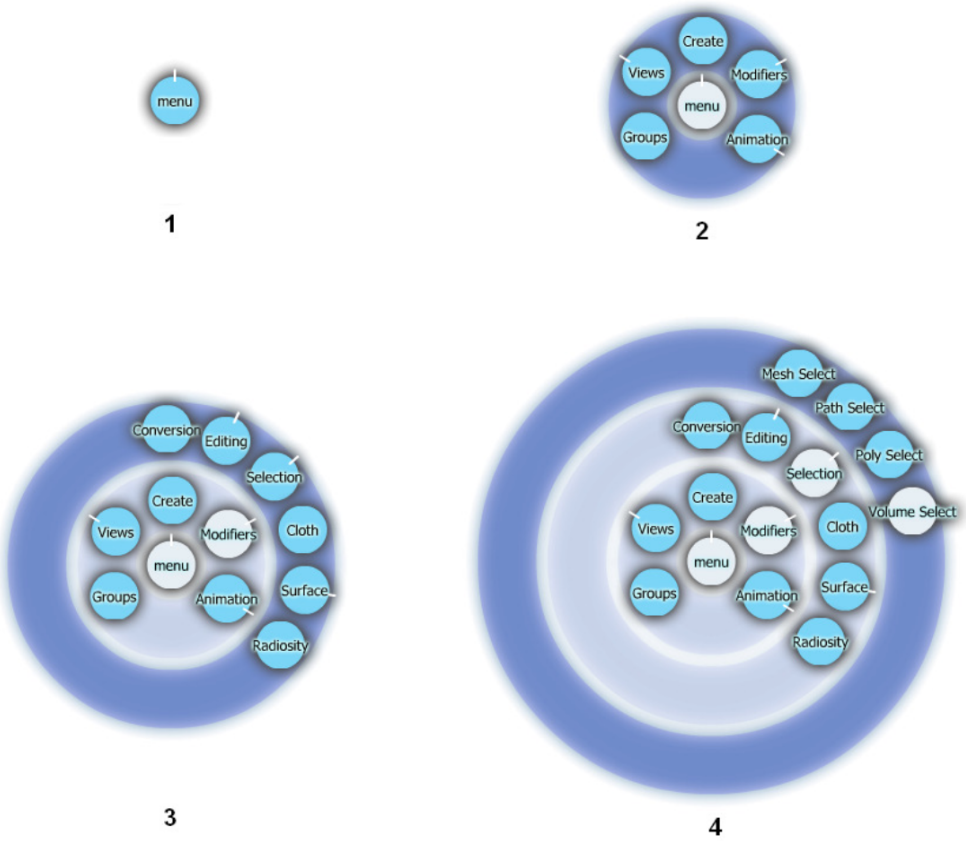
Figure 2.6: Different variants of existing menu interfaces (cont.)



(e) A list menu showing a filtered selection of options by utilizing a *searchbar*. The user typed in the query "grill", and receives options matched to his query. Picture taken from an article from the nngroup website[1].

[1] `https://www.nngroup.com/articles/site-search-suggestions/?lm=search-visible-and-simple&pt=article[Accessed:Nov.18,2024]`

Figure 2.6: Different variants of existing menu interfaces (cont.)



(f) An *(extending) radial menu* by Samp and Decker [2010]. The menu works with hierarchically structured options: When the user clicks on an option of the current hierarchical level (one of the circles on the outer ring), another ring appears, revealing further options at a lower level in the hierarchy.

Figure 2.6: Different variants of existing menu interfaces (cont.)

Chapter 3

Exploring Menu Interfaces for the Classification of Dark Patterns

As explained in Chapter 2.2.1, “The Dark Pattern Learning Game by Fiedler [2024]”, Fiedler [2024] only implemented high-level categories from Gray et al. [2023] for the game task of classifying dark patterns in his learning game. This game part was more disliked by players than the "finding" dark patterns part. The goal of our thesis is to improve on this aspect; we aim to integrate all 65 categories from the ontology by Gray et al. [2024] into the game. To do this, we explore viable tasks and user interfaces for the classifying part of the game in this chapter. We first design prototypes for menu interfaces in Chapter 3.1, “Prototyping”. Then, we examine notable differences in these prototypes to divide design considerations into selection tasks, interface styles and category definition representations in Chapter 3.2, “Tasks, Styles and Definition Representations”. There, we also present our final implementations and further additions connected to the classifying task.

3.1 Prototyping

We started our research by creating paper prototypes for the menu interfaces. In the following, we present the original version by Fiedler [2024], and four of our prototypes that have played a significant role for the final implementations of menu interfaces later. Additional paper prototypes can be found in Appendix A, “Paper Prototypes”. Whenever we use an *italic* font in this chapter, we refer to a type of menu defined in Chapter 2.3, “Viable Designs for Menu Interfaces”.

The original version only included high-level categories.

The original menu interface for selecting dark pattern categories by Fiedler [2024] is a *linear menu* interface. It only includes high-level categories to select, which are the same for every specific dark pattern on the page. See Figure 3.1 for a screenshot of the menu.

Figure 3.1: The original menu interface for selecting dark pattern categories by Fiedler [2024]. It is a *linear menu* 2.3, “Viable Designs for Menu Interfaces” and only includes high-level categories. The user can also hover over the "i" icons on the right side of the buttons to access tooltips, that present a definition for the respective category.

The first idea we had to extend Fiedler's design was to change the task of selecting high-level categories to selecting categories on the low-level (or meso-level, if there are no further low-levels associated with it). For this, we started with a simple *linear, sequential* menu. When the user clicks on a high-level category, the underlying hierarchy level of meso-level categories belonging to it replaces the current options. Analogously, when the user clicks on a meso-level category that has low-level categories attached to it, these are presented, where the user then can make their final selection of a low-level category. See Figure 3.2 for a photo of the prototype.

Our first ideas for extending the original design include *linear* and *radial* sequential menus, that make the selection of low- and meso-level categories possible.

Next, we designed a *radial, sequential menu*. This employs the same, sequential structure of choice like the first idea, but in a radial design. The options are in circular forms and organized in a circle. The category names are referred to by allegorical graphics and revealed when hovering over the options. When the user clicks an option, the underlying hierarchy level of categories gets revealed around it, replacing the current options. Again, the final selection is either a meso-level or low-level category. See Figure 3.3 for a photo of the prototype.

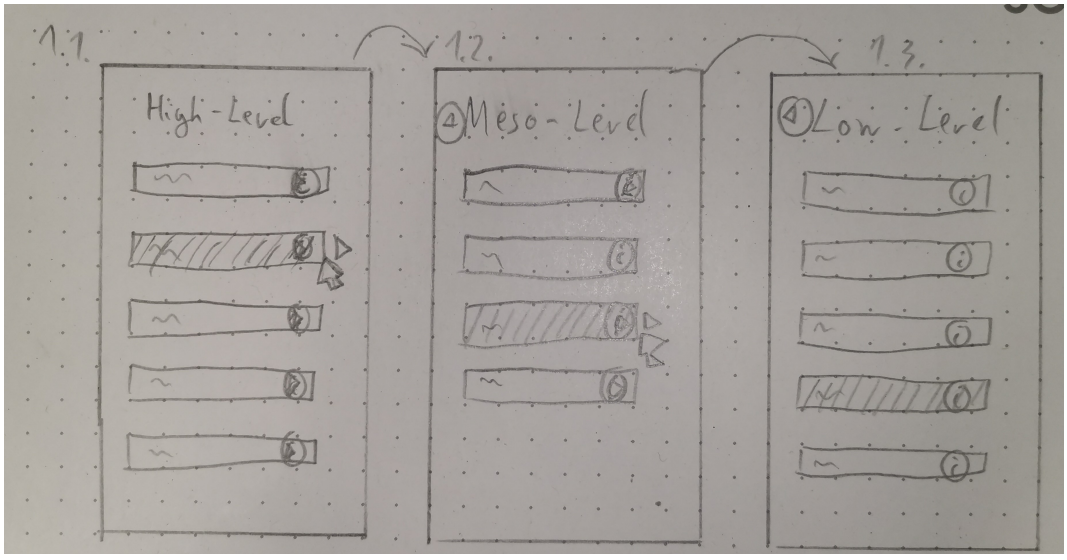


Figure 3.2: The first paper prototype: A *linear, sequential* menu interface. Clicking on an option reveals the underlying hierarchy level of categories, replacing the current options.

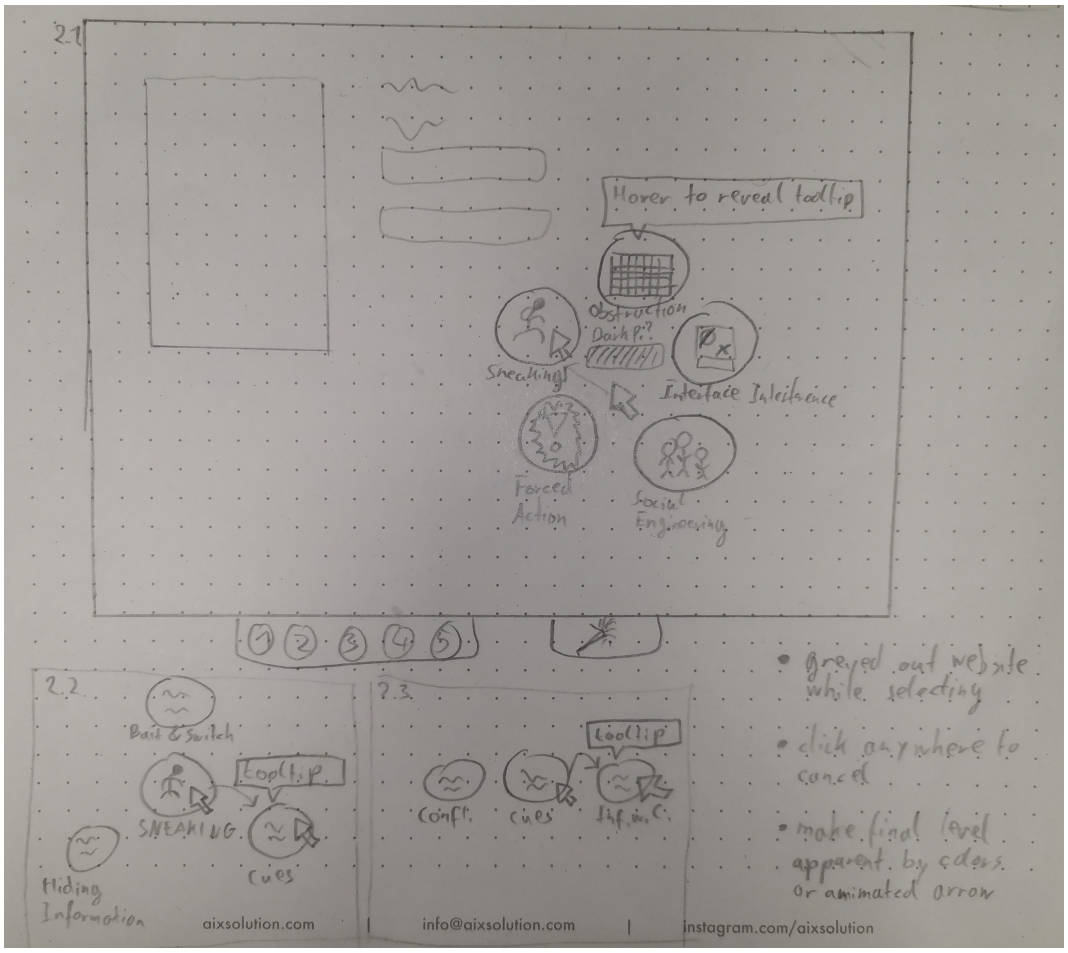


Figure 3.3: The second paper prototype: A *radial, sequential* menu interface. Clicking on an option reveals the underlying hierarchy level of categories around it, replacing the current options. Category names are referred to by allegorical graphics and revealed when hovering over the options.

The third paper prototype employs a reversed selection structure based on subsets of categories.

We realized, that, as we are designing menu interfaces for a game task, we are not restricted to always cover the whole ontology of 65 dark pattern categories for each selection. This gave us the idea of quiz-like subsets of categories as selections. Instead of going down the hierarchical path from high- to low-level categories, we could start at the lowest level, and directly present the player with a curated choice of low-levels (or meso-levels) to select. The third paper prototype aims to visualize this idea, and adds another concept: instead of making a final selection at the lowest level, the game could prompt the user to further select the ac-

cording higher meso-level and high-level categories from additional subsets. By expanding the selection task in this way, the learning of the hierarchical relations of the categories could be retained. See Figure 3.4 for a photo of the prototype. Besides these structural changes, it employs the same *radial, sequential* design as the second prototype.

The fourth paper prototype employs the metaphor of an encyclopedia, and evolved from a menu interface to a place for category definitions.

For the next prototype, we aimed to convey the metaphor of an encyclopedia. We quickly realized, that this design was much better suited for the presentation of category definitions than for a menu. Initially, definitions were implemented by Fiedler [2024] with tooltips the user could access in the original selection (see Figure 3.1). Our design offers the user a digital "book" to look up definitions and exemplary pictures for categories during gameplay. It can be traversed hierarchically in the means of the high-to-low structured ontology. See Figure 3.5 for a digital version of this prototype. We also include the original paper prototype in Figure A.2.

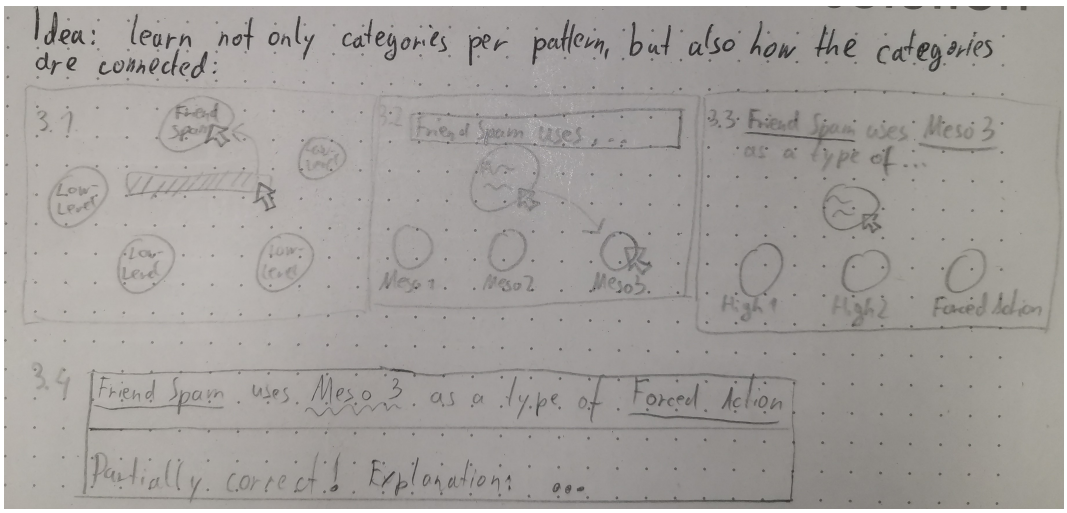


Figure 3.4: The third paper prototype: A *radial, sequential* menu interface based on curated subsets of categories. It employs a reverse order selection structure (from low-levels to high-levels), prompting the user to select not only the lowest-level category, but also the according higher-level categories. The intended learning effect of retaining the hierarchical relations of the categories is implied with the "result" box at the bottom, which tells the user that they have selected the low- and high-levels correctly, but not the meso-level, in this example.

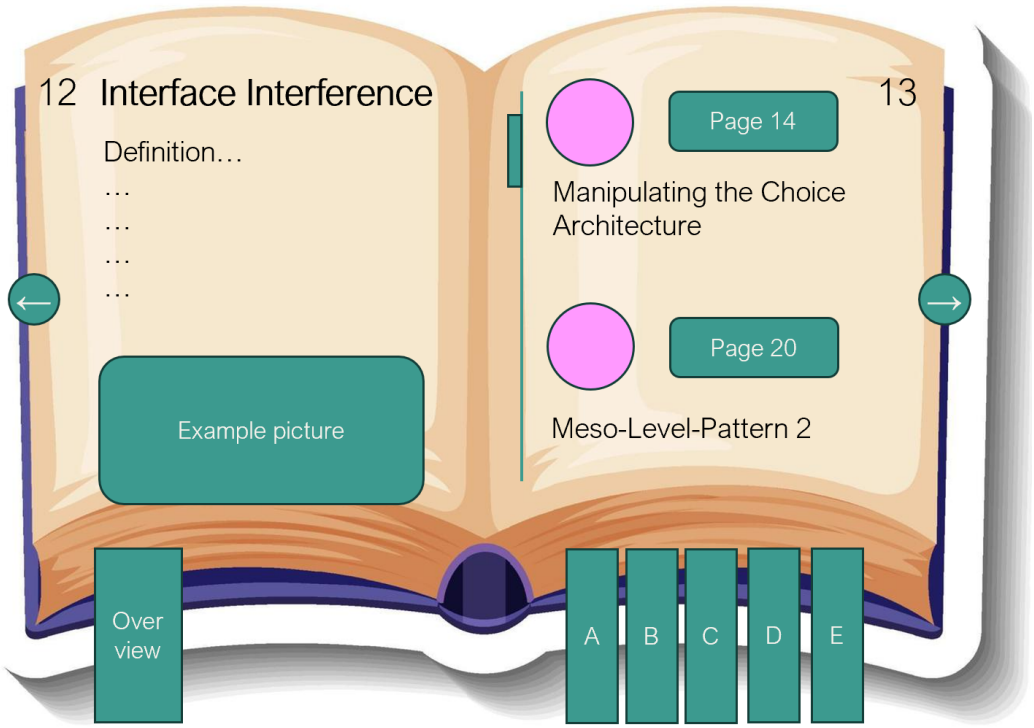


Figure 3.5: A digital version of the paper prototype employing the metaphor of a dark pattern encyclopedia. It offers the user a "book" to look up definitions and exemplary pictures for categories during gameplay. For every dark pattern category, the left page contains the name, definition and an example picture of it, and the right page a list of options for categories from the underlying level. The options can be clicked to get redirected to the respective page for the lower-level category. Bookmarks at the bottom redirect directly to an overview page of high-level patterns, or to one of the 5 high-level patterns, respectively. Background graphic created by brgfx[1].

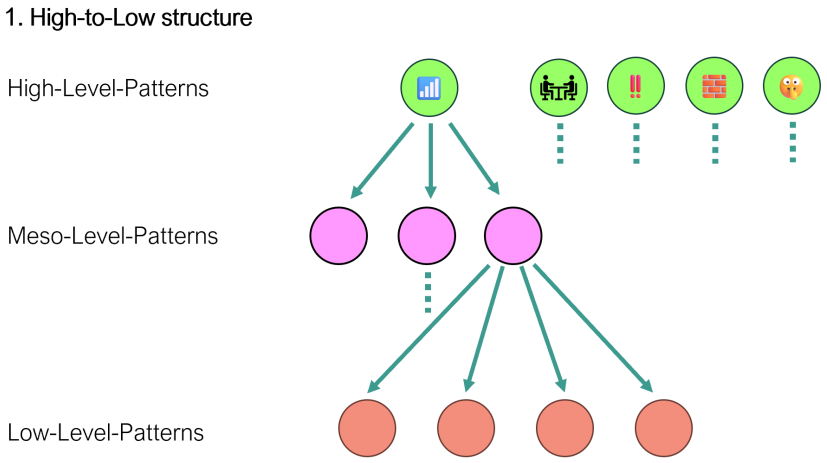
[1] https://www.freepik.com/free-vector/blank-book-white-background_3975643.htm[Accessed:Nov.19,2024]

3.2 Tasks, Styles and Definition Representations

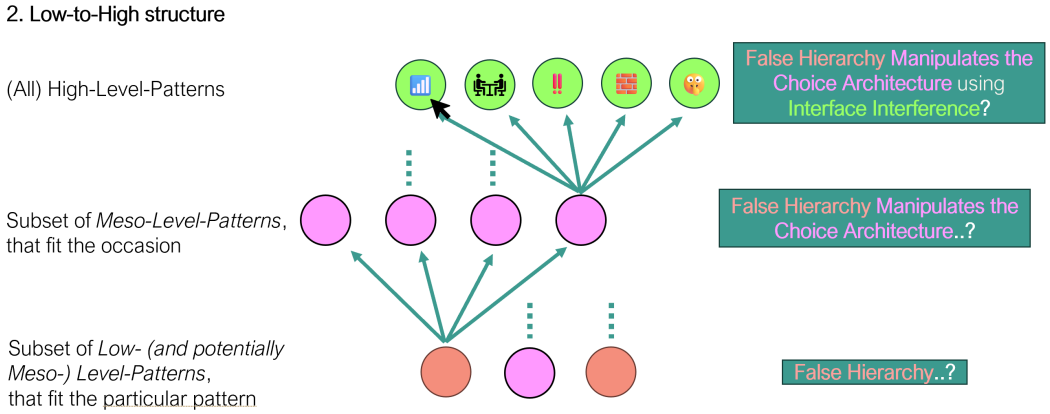
When we examine the similarities and differences of the paper prototypes we introduced, we can divide their design considerations into different, isolated aspects. Prototypes 1 & 2 (see Figures 3.2 and 3.3) essentially employ a high- to low-level structured *selection task*, in which the user has to select the lowest levels in the hierarchy. The differences between the two is a different *visual style* of the menu interface, namely *linear sequential* and *radial sequential*. Prototype 3 (see Figure 3.4) utilizes a *radial sequential* menu again, but differs in the prompted selection task: the user is required to make a low- to high-level structured selection from subsets, having to select multiple categories from different hierarchy levels for the same dark pattern. Now, we can identify two different aspects for our designs.

We identify two different aspects for our designs: selection tasks, and interface styles.

First, the *selection task*, being what category or categories, on which hierarchical levels, and from which solution sets (e.g. a curated subset or the full ontology), the user has to select. In our paper prototypes, these were the *high-to-low* and *low-to-high* selection structures (see Figures 3.6a and 3.6b for visualizations). However, we can also reduce these two to the structures of *only-high*, which is the original selection task of only selecting a high-level category, and *only-low*, where the user has to only select a low-level from one subset of categories per dark pattern.
The second aspect we identify is the visual *style* of the interface, in our case either linear sequential or radial sequential. For our purposes, we rename these styles to the *sidebar UI* and the *bubbles UI* styles respectively, as the linear menu is always in the sidebar at the right of the game page, and our implementation of the radial menu options looks similar to bubbles.



(a) The *high-to-low* selection structure. The user has to traverse the hierarchy starting from the high-levels, until they get to the lowest levels, where they make their final selection. Note that the final selection may also be a meso-level category that has no underlying low-levels.



(b) The *low-to-high* selection structure. The user has to traverse the hierarchy starting from a subset of curated lowest levels (low- or meso-levels with no underlying low-levels), where they have to make their first selection. Afterwards, they get presented with another curated subset of categories from the hierarchy level above their first selection, where they have to make their second selection. If their first selection was a low-level category, they get a final, third curated subset of categories from the hierarchy level above their second selection, where they have to make their final, third selection. As the user selected two or three categories from different subsets, the selections can be graded individually by the game.

Figure 3.6: A visualization of the *high-to-low* and *low-to-high* selection structures for the classification of dark patterns as a game task. The selection structures are independent of concrete menu interface implementations like "linear" or "radial" UIs.

Notably, these two aspects can be combined freely and are not limited to the combinations of "high-to-low/linear", "high-to-low/radial", and "low-to-high/radial", as employed by our paper prototypes. As we propose 4 selection tasks and 2 styles, we get a total of 8 possible menu interfaces to explore further, which we do in Chapter 3.2.1, “Selection Tasks & Interface Styles”, ordered by the selection tasks.

We can combine the found selection tasks and interface styles freely to a total of 8 menu interfaces to explore.

Furthermore, prototype 4 (see Figure 3.5) employs the metaphor of an encyclopedia for dark pattern category definitions. We further explore the idea in Chapter 3.2.2, “Category Definition Representations”, where we compare it to a modified version of the original representation of category definitions by Fiedler [2024], the "tooltips".

We further explore the metaphor of a dark pattern encyclopedia in comparison to the original tooltips.

3.2.1 Selection Tasks & Interface Styles

In this chapter, we will introduce the implementations of the 8 identified menu interfaces, ordered by their 4 selection tasks. For each selection task, the interface is implemented with the *sidebar UI* and the *bubbles UI* styles. We also propose the *searchbar* menu interface, being a unique form of the "only-low" selection task.

General design considerations and changes to the original menu interface by Fiedler [2024] include:

General design changes include adaptive background colors, emojis, and considerations for the *bubbles* style UIs.

- The background tint of buttons adapts to the hierarchy levels of the categories: purple for high-levels, blue for meso-levels, and green for low-levels. This color-scheme is established in the whole game, at every situation where the hierarchy levels are mentioned, for example in our new feedback screens that are introduced later in this chapter.
- We identify a representing emoji for each of the 65 dark pattern categories. The emojis are used as addition to the category names in *sidebar* styled interfaces, and as substitute for the names in *bubbles* styled ones.

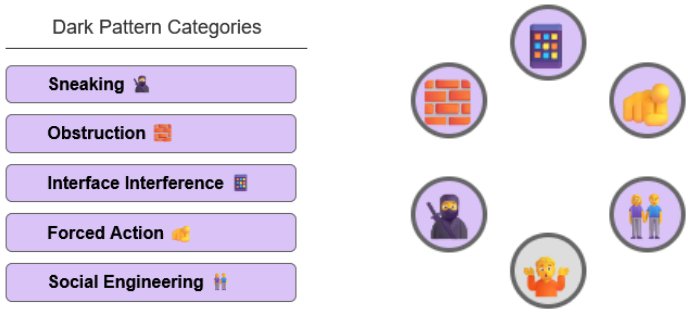
- A right-pointing arrow in a button represents, that the button does not perform a final selection, but presents another selection of options from a connected hierarchy level.
- For the *sidebar* UI style, if the list of options gets too long, it is cut off and made accessible via a scrollbar.
- Menu interfaces in the *bubbles* UI style appear at the place the user has clicked a potential dark pattern element. The options surround the mouse cursor in a circle. The "back" and "I'm not sure" buttons, that are found at the bottom of *sidebar* styled UIs, are always in the middle and bottom of the circle, respectively. When clicking anywhere on the page, where there is no other dark pattern element, the interface disappears.
- For the *bubbles* UI style, a categories name is presented in a little box, that is visible when the user is hovering over a button.
- For the *bubbles* UI style, components like the timer, score display or the Magic Wand button, that were contained by the sidebar, move to the bottom of the screen, so that the sidebar can be hidden completely.

The *Only-High* selection task is structurally similar to the original.

The *Only-High* selection task (see Figure 3.7) is almost identical to the original classifying game task by Fiedler [2024]. Differences include the change of high-levels from the updated ontology by Gray et al. [2024] and the general considerations from above. The user has to only select the correct high-level category.

The *High-To-Low* selection task follows the hierarchy starting from the top.

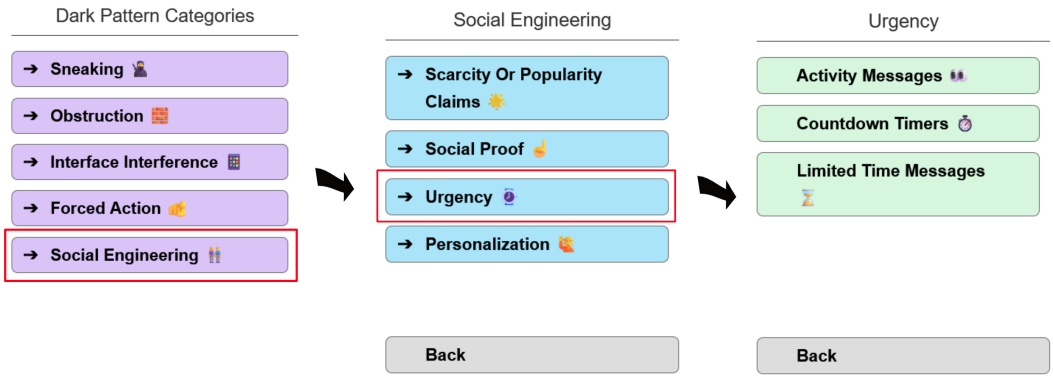
In the *High-To-Low* selection task (see Figure 3.8), the user has to only select the correct lowest level category, which may be a low-level or a meso-level category with no underlying low-levels. To find the lowest level category, the user starts at the top of the hierarchy with the high-levels. Then, they can move down the hierarchy: by choosing a category with an arrow, the categories from the underlying hierarchy level get presented instead of the current options. The user can also always go back up in the hierarchy by clicking the "Back" button.



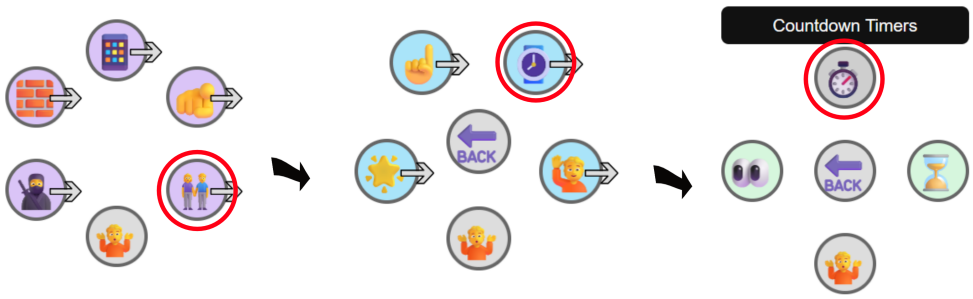
(a) The *sidebar* style *only-high* menu interface, similar to the original by Fiedler [2024].

(b) The *bubbles* style *only-high* menu interface.

Figure 3.7: The *only-high* menu interfaces, including all 5 high-level categories.



(a) The *sidebar* style *high-to-low* menu interface.



(b) The *bubbles* style *high-to-low* menu interface.

Figure 3.8: The *high-to-low* menu interfaces. The user has to only select the correct lowest level category, which they can find by following the hierarchy from high-level to low-level. In this example, the path of options the user selects is outlined in red.

The *Subsets of Only-Low* selection task prompts the user to select a lowest level category from a curated subset.

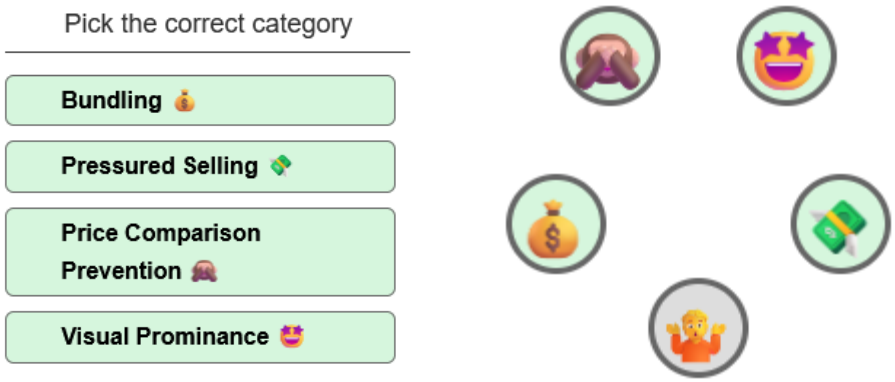
In the *Subsets of Only-Low* selection task (see Figure 3.9), the user has to only select the correct lowest level category, which may be a low-level or a meso-level category with no underlying low-levels, from a curated subset of solutions. These subsets are created individually for each dark pattern category, with the aim of providing a reasonably difficult challenge in working out the correct solution. For example, a dark pattern element that employs "Countdown Timers" may have the subset selection "Countdown Timers", "Limited Time Messages", and "Confirmshaming", which superficially may appear as similar categories. Note, that subsets always have to include the correct answer.

Elements that are actually no dark patterns, and elements that include multiple correct categories are accounted for.

Further, elements that are actually no dark patterns, and elements that include multiple correct categories from the same hierarchy level are accounted for. If an element is not a dark pattern and has no category preset, a random, lowest level category gets selected at the start of the level, so that the corresponding subset is displayed and the user can't deduce the element not being a dark pattern simply by looking at the selection of categories. We also added the functionality of presetting a "false" category to a non dark pattern element, of which the subset then gets presented instead. This allows to preset intentionally difficult choices for the user deciding on an element being a dark pattern. Finally, for elements that have multiple correct solutions, a mixed subset is created containing all correct solutions and an evenly distributed amount of options from the corresponding, individual subsets.

The *Subsets of Low-To-High* selection task follows the hierarchy starting from the bottom.

The *Subsets of Low-To-High* selection task (see Figure 3.10) extends the Subsets of Only-Low selection task and is arguably the most complicated one, as the user has to determine multiple solutions. First, they have to select the correct lowest level category, which may be a low-level or a meso-level category with no underlying low-levels, from a curated subset of options. Then, they have to select the correct higher level category that **corresponds to the lowest level category** they selected first, from another subset of options. For example, a question for the second subset may be "To which higher level category does Countdown Timers correspond?" and *not* "Which meso-level category is this particular dark pattern element?". This is an impor-



(a) The *sidebar* style *subsets of only-low* menu interface.

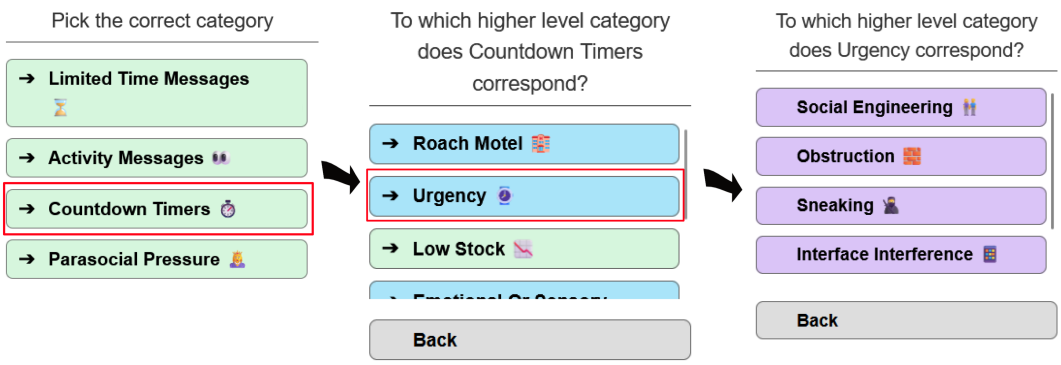
(b) The *bubbles* style *subsets of only-low* menu interface.

Figure 3.9: The *subsets of only-low* menu interfaces. The user has to only select the correct lowest level category from a curated subset of solutions.

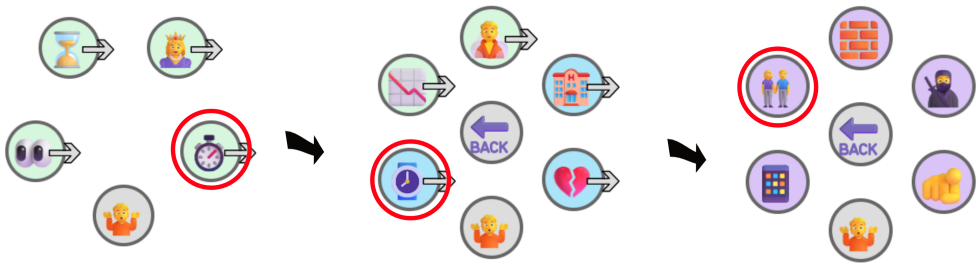
tant difference, as this selection task aims prompts for the lowest level category and in addition, its relation in the hierarchy, instead of asking the user individually for multiple categories related to the dark pattern element. If the lowest level category was a low-level one, the user further has to select the corresponding category from a third subset of higher level categories, which in that case is always a high-level one.

The higher-level subsets depend on the previously selected lower-level categories.

Elements that are actually no dark patterns, and elements that include multiple correct categories from the same hierarchy level are accounted for in the same way as described in *Subsets of Only-Low*. Note, that subsets always have to include the correct answer. Further, the second (and third) subset depends on the selection made before. For example, if the user selected "Countdown Timers" at first, the next selection will be tailored to a set of the corresponding higher level category, in this case "Urgency". If the user chose the first category incorrectly, they get presented the corresponding higher-level set for that incorrect category. This aims to reduce the chance of the user figuring out the solution of the first subset by inferring from following subsets. However, it also requires for a re-design of the score evaluation and feedback overlay, as to not confuse the user.



(a) The *sidebar* style *subsets of low-to-high* menu interface.



(b) The *bubbles* style *subsets of low-to-high* menu interface.

Figure 3.10: The *subsets of low-to-high* menu interfaces. The user has to select multiple correct categories from respective subsets consisting of options from either two or three respective hierarchy levels. The selection is made hierarchically, starting from the subset including the correct lowest level category of the concrete dark pattern element. The following subsets are dependent on the first selection. The amount of solutions required depends on the lowest level being meso- or low-level. The user can also always go back down in the hierarchy by clicking the "Back" button, and change their first selection, if wanted. In this example, the path of options the user selects is outlined in red.

The feedback and explanations overlay has been re-designed.

To account for the multitude of solutions required by *Subsets of Low-To-High*, we re-designed the feedback and explanations screen by Fiedler [2024], extending their implementation. Now, there are five possible situations, in which different versions of the screen are presented:

1. The selection is completely correct (see Figure 3.11a). The selected category or categories are presented with a green check symbol to indicate that they are correct. These may be: one high-level category; one lowest level category; a low-level, a meso-level and a high-level category; or a meso-level and a high-level category, depending on the selection task and specific element. Further, for any of these cases, the complete correct classification is presented, including the correct categories from every applying hierarchy level. Optionally, a preset explanation text for the specific element is included. The user gets the full score.

2. The selection is incorrect (see Figure 3.11b). For Subset of Low-To-High, this applies as soon as the first selection of the lowest level category is incorrect. The selected category is presented with a red cross symbol to indicate that it is incorrect. Again, the complete correct classification is presented, and optionally, a preset explanation text is included. The user gets a neutral score, as they still found a dark pattern. They can also choose to explain in a text input field, why they thought the dark pattern was from a different category.

3. The selection is partially correct (see Figure 3.11c). This applies only for the Subset of Low-To-High selection task, and only if the first selection is correct, but one (or two) of the later selections are incorrect. All correct categories from every applying hierarchy level are presented in a comparison with the selected (two or three) categories, including a green check symbol if the selected category was correct, and a red cross symbol if it was not. Optionally, a preset explanation text is included. The user gets half of the score, as they still classified the dark pattern correctly on the lowest level.

4. The user selected "I'm not sure". The selection "Unsure" is presented with a red cross symbol. The complete correct classification is presented below it, and optionally, a preset explanation text is included. The user gets a half of the score, as they still found a dark pattern and didn't answer incorrectly.

5. The element actually was not a dark pattern. The user gets notified about this and loses some of their score, as they incorrectly assumed a dark pattern where there was none. They can also choose to explain in a text input field, why they thought the element was manipulative.

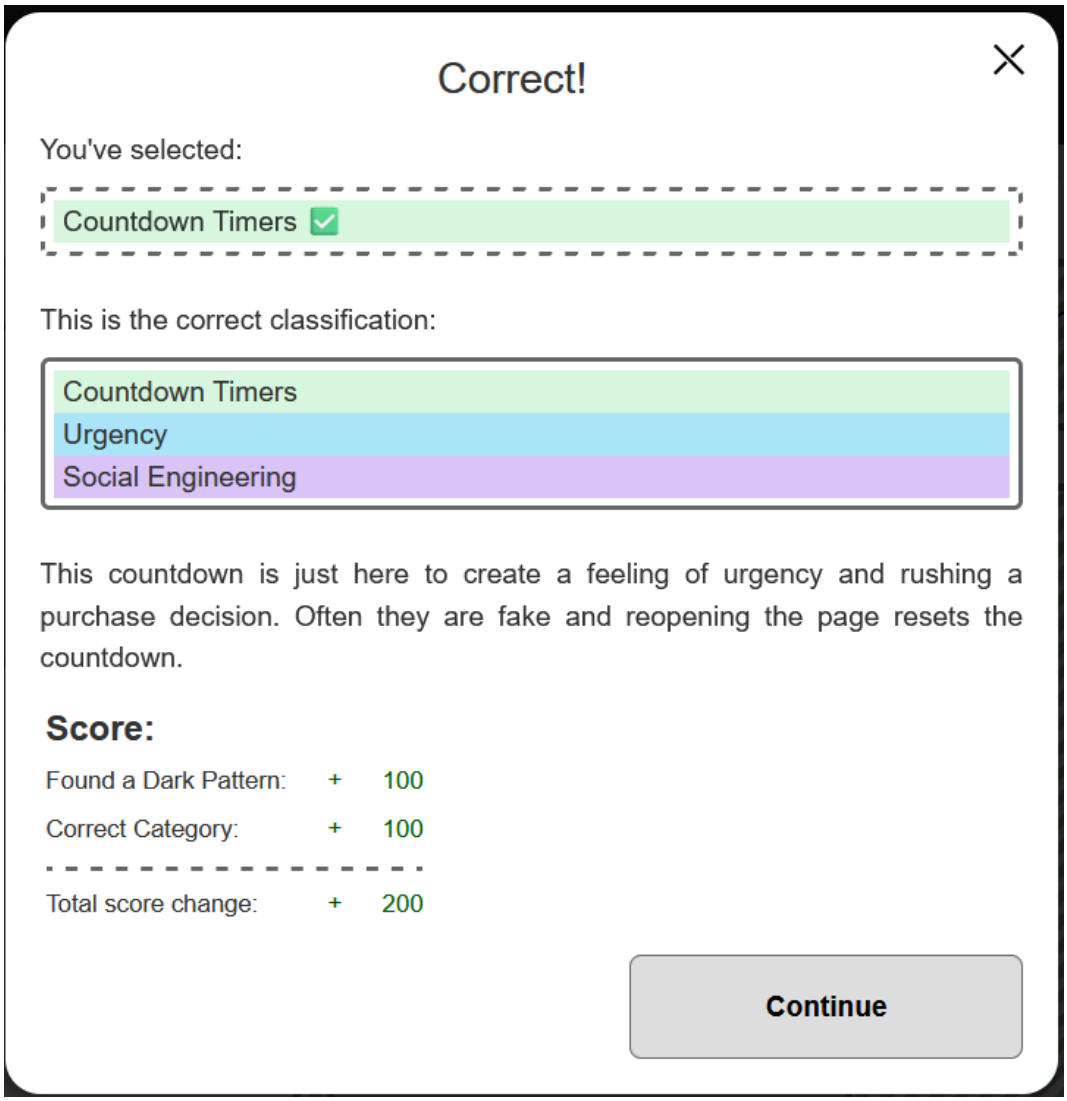
The *Searchbar* menu interface is a special form of an *Only-Low* task in the *sidebar* UI style.

The *Searchbar* menu interface (see Figure 3.12) prompts the user to select a category from **all** lowest level categories in the ontology. It is implemented only in the sidebar and has no corresponding concept in the *bubbles* UI style. The user gets presented a full list of all lowest level categories, accessible by a scrollbar. The list is ordered by applying depth-first search on the hierarchy. Now, the user can filter the list by typing into the input field above the list. The input string is matched to the category names. We chose to make the selection task a type of *Only-Low* with no subsets, as the use of the searchbar couldn't be explored adequately with the other selection tasks above. For example, filtering a list of only five high-level elements may not be of much help, and extending the selection task to also prompt higher-level categories would have biased the study focus away from the sole effectiveness of the searchbar.

We payed attention to preserve all existing game mechanics by Fiedler [2024] for every new menu interface.

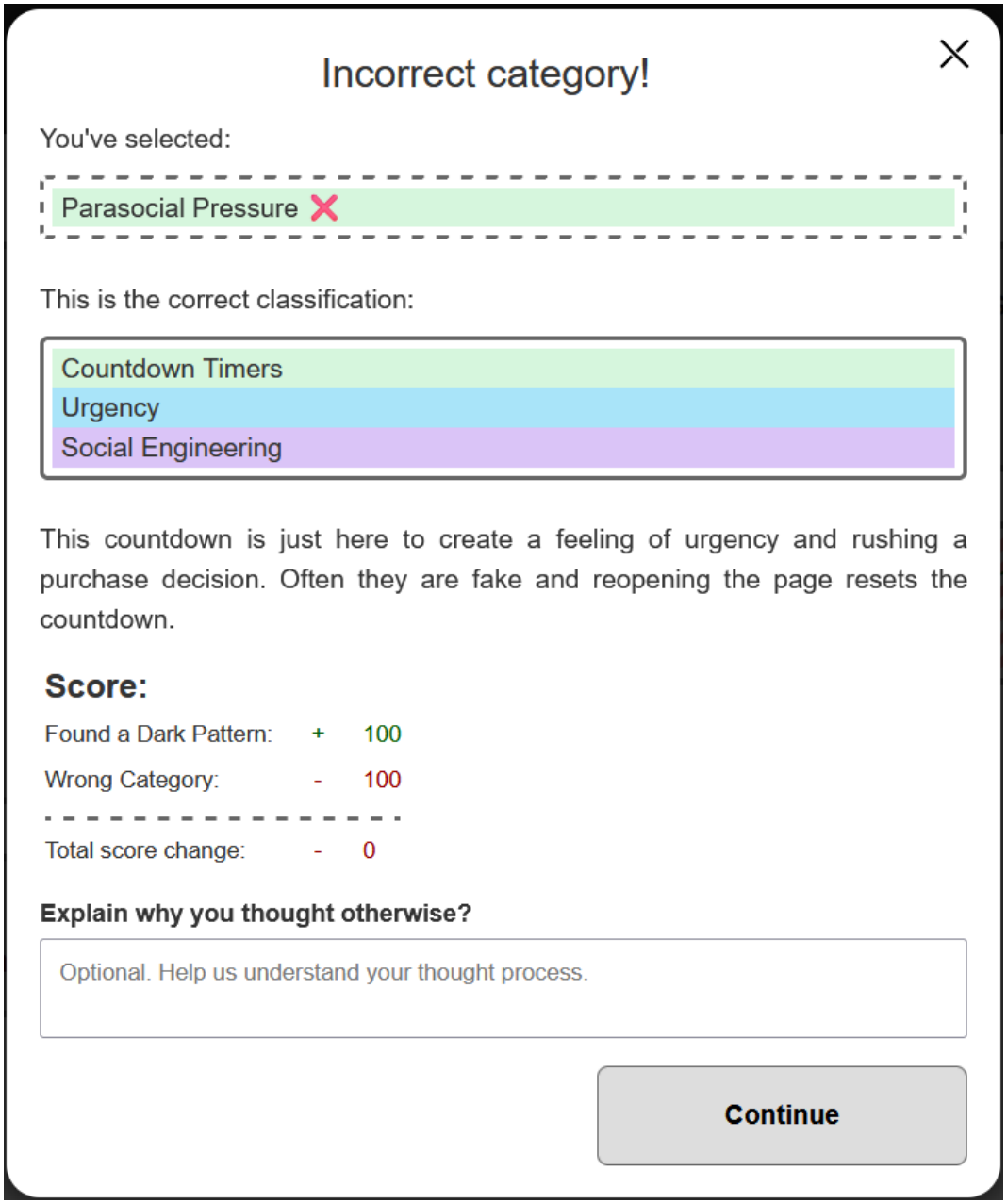
As the new selection tasks prompt for lowest-level categories, the dark pattern elements in the game levels have to be updated to fully employ the new functionalities. In Fiedler [2024], these elements were referenced with a simple attribute, for example "dpSneaking". Analogously, it is sufficient to add a lowest dark pattern category name to an element as a correct solution, e.g., "dpCountdownTimers" automatically includes the higher level categories "dpUrgency" and "dpSocialEngineering" for the solution set of an element. Our final implementations also work with all former game mechanics and levels employed by Fiedler [2024]: For example, dark pattern elements with

no specified low- or meso-level categories have their preset high-level correct category interpreted as the lowest one by the new selection tasks. This is possible because of a new, sophisticated data structure we designed, conveying the dark pattern ontology and its relations. It has been integrated in all legacy systems that involve dark pattern categories.



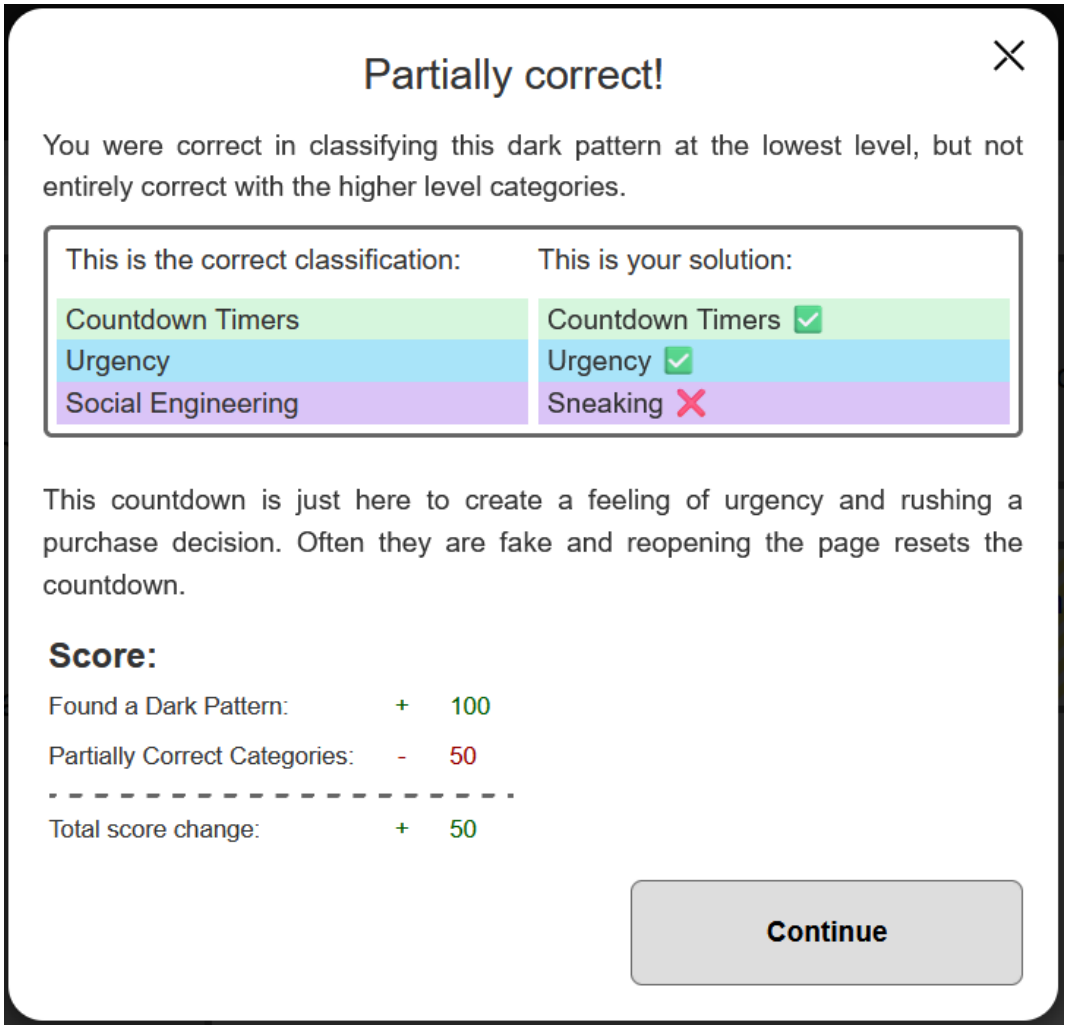
(a) The feedback screen for the case, that the selection is completely correct. The selection task prompted only the selection of the low-level category.

Figure 3.11: Three variants of the re-designed feedback and explanation overlay



(b) The feedback screen for the case, that the selection is incorrect. Depending on the selection task, the user may have been prompted to only select the low-level category, or to select the low-level category and then its higher levels. In either case, the user failed at the first selection.

Figure 3.11: Three variants of the re-designed feedback and explanation overlay (cont.)



(c) The feedback screen for the case, that the selection is partially correct. The selection task prompted the selection of the low-level category and then its higher levels. The user failed to correctly select the high-level category corresponding to "Urgency", for this example element.

Figure 3.11: Three variants of the re-designed feedback and explanation overlay (cont.)

3.2.2 Category Definition Representations

The ontology by Gray et al. [2024] features precise definitions for each dark pattern category. In this chapter, we examine two methods to represent these definitions: the *Tooltips*, a slightly modified version of the representation by Fiedler [2024]; and the *Book Of Dark Patterns* introduced before in Chapter 3.1.

The *Tooltips* are directly accessible during the selection of categories in sidebar style UIs.

Our *Tooltips* (see Figure 3.13) are a modified version of those by Fiedler [2024]. They can be accessed only while selecting a category in the *sidebar* style UIs, as there was no sensible way of integrating them with the reduced *bubbles* UI style. Instead of originally being presented as pop-ups above the selection options, we reserved place for them at the bottom of the sidebar. When the user hovers over an "i" icon in one of the options, the according definition appears below.

The *Book Of Dark Patterns* offers definitions, examples, and relations of all categories at any time.

The *Book Of Dark Patterns* (see Figure 3.14) offers the user a way of looking up information on all dark pattern categories from the ontology. The book is available via a button (either in the sidebar or at the bottom of the page) at any time during the whole gameplay, instead of only during the classifying task. It is presented as an overlay, filling a large part of the screen. The traversal process follows the high-to-low structured hierarchical ontology, starting at a selection of the five high-level categories. For every category, the left page contains the name, definition and an example picture of it, and the right page a list of options for categories from the underlying level, if existing. The options can be clicked to get redirected to the respective page for the lower-level category. We chose to not include the bookmarks from the prototype, as we worried the interface may get too cluttered with buttons. The book can be closed by clicking anywhere outside of it on the page.

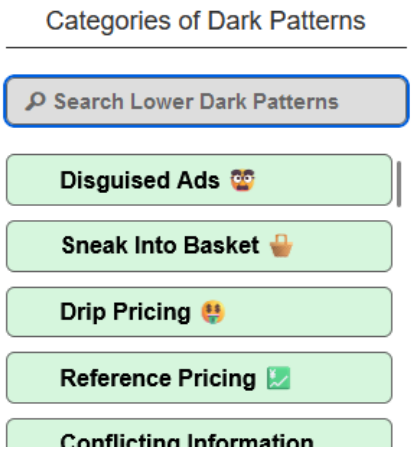


Figure 3.12: The *Searchbar* menu interface is a special form of an Only-Low task in the sidebar UI style. The user gets presented a full list of all low- and meso-level categories, that can be filtered by typing into the input field above it.

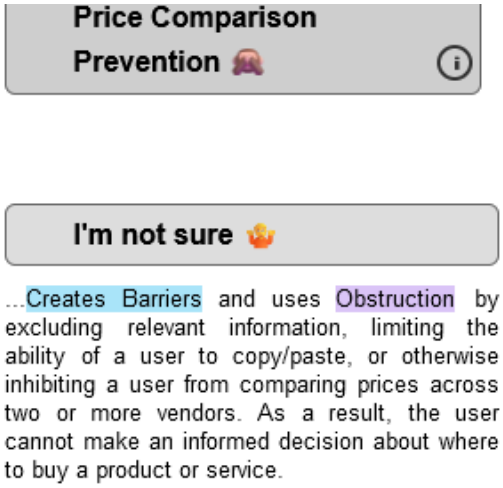


Figure 3.13: Our modified version of the *Tooltips*, a category definitions representation. They are only available while selecting categories in the sidebar and appear at the bottom of it, when the user hovers over an "i" icon in one of the options. Mentions of category names in the definitions are colored by their respective hierarchy level.

Figure 3.14: The *Book Of Dark Patterns*, a category definitions representation. The book offers the user a way of looking up information during the whole gameplay. For every dark pattern category, the left page contains the name, definition and an example picture of it, and the right page a list of options for categories from the underlying level. The starting page presents the user with the choice of one of the five high-level patterns (top left). After deciding on a high-level category, the user is redirected to the selected categories page and presented a selection of the corresponding meso-level categories (top right). After choosing one of these options, the user gets further redirected to the selected categories page and the corresponding low-level categories (bottom left). Finally, the user can select a low-level category to view its information (bottom right). Here, the user can also click the colored meso- and high-level category names to get redirected directly to their corresponding pages. The arrow buttons on the left and right side of the book can be used to switch to pages of parallel categories, that belong to the same higher level category as the current page one.

Chapter 4

User Study, Evaluation & Interpretation

In this chapter, we describe the user study we conducted to evaluate our menu interfaces, that we developed in Chapter 3, "Exploring Menu Interfaces for the Classification of Dark Patterns". In Chapter 4.1, "Methodology", we describe the study design and goal. In Chapter 4.2, "Results & Evaluation", we present our findings and evaluate them.

4.1 Methodology

We aim to evaluate the effectiveness of our menu interfaces with a user study.

We designed a user study with two goals in mind: Firstly, we want to evaluate which of the menu interfaces we designed are best suited for the classification game task. Secondly, we want to evaluate, if our new menu interfaces can yield a better understanding of the workings and harms of specific dark patterns. We assume, that the inclusion of the ontology in our interfaces may achieve this. Our assumptions are grounded on the revision of Bloom's taxonomy by Krathwohl [2002]. We want to explore if including all dark pattern categories and their hierarchical connections leads to a higher cognitive process dimension in the conceptual knowledge dimension.

We are particularly interested in the change to a higher cognitive process dimension due to our contributions.

In particular, we are interested in the change from merely *recognizing* (remembering) dark patterns to *classifying* (understanding) them during the classifying game task. Note, that Fiedler [2024] assumes the original game to be already on an "evaluate" cognitive process level. However, as we focus on the classifying game task in our research, we argue the specific level of that task to be only the "remembering" level in the original implementation, as players often didn't understand why a dark pattern is classified in a certain high-level category. Finally, we hope that a higher cognitive process dimension for the classifying game task also leads to a beneficial learning effect for the overall recognition of and confidence dealing with dark patterns.

We derive the following research questions:

RQ-1 Which menu interfaces are most preferred?

RQ-2 Which menu interfaces allow for the quickest selection?

RQ-3 Do players better understand specific dark patterns after classifying them with the menu interfaces?

4.1.1 Study Design

For each participant, our within-subjects user study covered every task, style and definition presentation from our menu interfaces. We counterbalanced by employing different orders of the menu interfaces.

The study was designed as a within-subjects experiment, where each participant played each selection ask and interface style, and encountered both the *Tooltips* and the *Book of Dark Patterns*. For this, each participant played 7 generated game levels in total. Each level employed a different combination of task, style, and definition presentation, covering all individual options (but not all combinations) by the end of the study. The first 4 levels employed the same style and definition presentation, and a different selection task (*Only-High, High-To-Low, Only-Low, or Low-To-High*) for each level. The order of the selection tasks was counterbalanced by using a Latin Square [Bradley, 1958], which resulted in 4 different orders. Further, half of the participants started with the *Bubbles* style + the *Book*; a quarter with the *Sidebar* and the *Book*; and a quarter with the *Sidebar* and the *Tooltips*.

The 5th and 6th level repeated the selection task the participant played last, but changed the style from *Sidebar* to *Bubbles*, and replaced the *Tooltips* with the *Book*, or vice versa, respectively. The 7th level always employed the *Searchbar* task, and had either the *Tooltips* included for half of the participants, or the *Book* for the other half. At the start of each level, the employed selection task was explained, and any change of style or definition representation mentioned.

Before we conducted the actual study, we performed a pilot run to identify potential shortcomings. Most of all, we improved on the phrasing and structure of questionnaires, which we built into the game app.

Procedure

The study was conducted in person and was completely playable in a single browser window.

We conducted the study in person, with participants playing the game on a laptop, and the screen mirrored for our observation. We started by informing the users about the study procedure, letting them sign an informed consent form. From this point on, the whole study, including all questionnaires, was conducted on the computer in a browser screen, and could be walked through without any interference from us. There was no time limit involved in the study, which the participants were made aware of. If the participants did anything we took note of and wanted to clarify, we asked them for an explanation of their actions after the study. All data we gathered was associated with anonymous, random user id's.

The participants filled out a demographics questionnaire.

The participants filled out a questionnaire on demographics, which contained questions regarding gender, age, field of study, and highest achieved degree. Also, questions on the amount of hours spent online daily, and the perceived, own familiarity with dark patterns were included.

The participants played through a tutorial, which explained 6 lowest level categories and their higher level ones.

Next, the participants were presented with a tutorial section. This was partly similar to the one originally employed by Fiedler [2024], however, we changed the focus on the explanation of the dark pattern ontology by Gray et al. [2024], its different hierarchy levels, and their connections. The participants were taught 5 low-level and 1 meso-level dark pattern categories, namely: *Disguised Ads*, *False Hierarchy*, *Confirmshaming*, *Limited Time Messages*, *Countdown Timers*, and *Trick Questions*. We picked these, as they feature overlapping meso- and high-level categories, which we also defined, namely: *Bait And Switch*, *Manipulating Choice Architecture*, *Personalization*, *Urgency*, *Sneaking*, *Interface Interference*, and *Social Engineering*. We used example pictures to aid the definitions of the 6 lowest level categories, which can be found in B.1, “Image Assets”. Also, the tutorial featured a brief explanation of the game and the following study process, including a mention of the newly added *Book of Dark Patterns*, and that each level included exactly 4 dark pattern elements.

The levels were randomly generated, including 4 dark pattern elements from a pool of 6 categories.

Next, the participants played 4 consecutive levels. As mentioned, these levels focused on the comparison of the four selection tasks we introduced. The levels were completely empty besides the 4 dark pattern elements, because we wanted to omit the "finding" dark patterns game task, and let the participants focus on the classifying task. At the start of each level, these elements, and their order on the page, were randomly picked from a pool, consisting of the 6 lowest level categories that were explained in the tutorial. For each category, we designed the elements to be similar to the images from the tutorial, each having multiple variants, which we created by varying their texts and colors. During the whole study, we also logged timestamps for different actions, for example to determine how long the participants took between the selection of an element and its final categorization.

Questionnaires were included to ask the participants about their opinions on the tasks, styles, and definition representations.

After playing through the first 4 levels, the participants filled out a first questionnaire, in which they could state their opinions on the selection tasks, in the order they played them. For each task, a screenshot of it was included, either in the *Sidebar* or in the *Bubbles* style, depending on which style the participant encountered. Further, for each

task, ten 5-point Likert scale questions (from "strongly disagree" to "strongly agree") and four free-text input fields were included, regarding strengths, weaknesses, suggestions for improvements, and further comments. The questions and free-text fields were the same for each task. After the questionnaire, the participants continued to the next two levels.

After playing levels 5 and 6, the participants were asked about their opinions on the *Sidebar* style, the *Bubbles* style, the *Tooltips*, and the *Book*, in the order they played them, in a second questionnaire. The participants had encountered each at least once at this point. Again, corresponding screenshots were used. For the *Sidebar* and the *Bubbles* styles, ten 5-point Likert scale questions and five free-text input fields were included, each. Eleven different 5-point Likert scale questions and five input fields were included for the *Tooltips* and the *Book*, each. The questions and input fields were identical for the pairs, besides replacing the words "sidebar" for "bubbles", and "tooltips" for "book", and vice versa, respectively. The additional free text question was, which style or representation was preferred by the participants. After the questionnaire, the participants continued to the final level.

After playing the 7th level, the participants first were asked about their opinions on the *Searchbar* UI in a third questionnaire. Again, a corresponding screenshot was used, and ten 5-point Likert scale questions and four free-text input fields were included. Next, the participants were asked to rank the selection tasks they encountered on three scales: "Overall", "Difficulty of Task", and "Ease of Use". Besides *Only-High, High-To-Low, Only-Low, and Low-To-High*, the *Searchbar* was also included as a selection task in the comparison. Also, corresponding pictures in the correct styles were included to remind the participants about the tasks. Further, the participants had to choose a favorite (existing) combination of selection tasks, interface styles and category definition representations. Finally, two free-text input fields could be filled out to state own ideas the participants may have had for novel menu interfaces, and final comments on the study. With that questionnaire, the study ended on a "Thank You for participating" screen.

All questionnaires are included in Figure B.2. Their designs are inspired by the questionnaires of Fiedler [2024] and use their existing, technical implementation.

4.2 Results & Evaluation

This Chapter includes the analysis, discussion and interpretation of the data we aggregated from the logged data, the questionnaires, and the notes during our study. First, we present demographics in Chapter 4.2.1. Then, we present the quantitative data from the logs and questionnaires in Chapter 4.2.2, “Quantitative Analysis”. Further, we present the qualitative data from the free-text answers of the questionnaires, and our notes, in Chapter 4.2.3, “Qualitative Analysis”. Finally, we evaluate the data in Chapter 4.2.4, “Discussion & Interpretation”.

4.2.1 Demographics

We included 8 participants, who all had played the original game by Fiedler [2024] before.

Because of the limited scope of our thesis, we could only include 8 participants, who were all employed in technical domains. Further, all of our participants had played the original game by Fiedler [2024] before, which we hoped allowed them to concentrate on our added menu interfaces.

4 of the participants were female, and 4 were male. They were between 21 and 31 years old (M=24.88, SD=3.62). 7 participants had a background in computer science, and 1 in media informatics. 4 participants had a high school diploma as their highest academic degree, two had a Bachelor of Science, and two a Master of Science degree. The participants stated being online between 3 and 11 hours (M=6.63, SD=2.87) a day. 6 participants stated to have "experience with the classification of dark patterns", and 2 to have "done extensive research on dark patterns". The participants took between 50 and 75 minutes to finish the complete study, of which between 4 and 8 minutes were spent on the tutorial.

4.2.2 Quantitative Analysis

97% percent of all dark pattern elements in the study were correctly classified by the participants.

Firstly, we take a look at the classification rates of the dark patterns. There were 4 dark pattern elements per level, which resulted in 224 total elements throughout all participants and levels. Of these 224 elements, 218 were classified correctly (õ97%); 12 were classified "incorrectly" (on the lowest selection stage); 4 were classified only "partially" correct (in the Low-To-High selection type); 1 was missed, i.e. not spotted on the website; and 1 was not classified by choosing the "I'm not sure" button. Notably, the low level category *Confirmshaming* was incorrectly guessed as being the meso-level category *Trick Questions* 6 times, and *Trick Questions* was incorrectly guessed 3 times as being *Confirmshaming*, making up 75% of all incorrect guesses. We intentionally designed the elements of the two categories to be similar in their appearance, however, we payed attention to make them clearly differentiable, by following their ontology definitions. Also, *Countdown Timers* were confused with being *Limited Time Messages* two times by the same participant, which they reasoned as *Countdown Timers* technically including the notion of a *Limited Time Message*. Further, *Trick Questions* was classified as being the lower level category for *Social Engineering* multiple times, instead of the correct answer, *Interface Interference*. The rest of incorrect, missed and partially correct selections could be largely attributed to the participants choosing options too quickly, realizing directly after, that they wanted to pick a different option.

We measured the timings for the categorization of elements, depending on their *categories*.

We also measured the time it took the participants to categorize an element after selecting it, in seconds. *Confirmshaming* and *Trick Questions* took the longest (M=17.91, SD=18.21 and M=15.74, SD=17.06), which supports our observations mentioned before. *Countdown Timers* were categorized the quickest (M=9.10, SD=6.42). Generally, we observe that the participants were quite different in their categorizing speeds, as indicates the rather large overall standard deviation of 12.37. We include all results about the accuracy and timings in classifying the dark patterns in Table 4.1, ordered by their lowest level categories.

	CR	IR	PC	US	MS	M	SD
Disguised Ads	31	1	1	0	0	10.44	7.13
False Hierarchy	36	0	0	0	0	13.32	15.8
Trick Questions	34	3	2	0	0	15.74	17.06
Confirmshaming	33	6	0	1	0	17.91	18.21
Limited Time Messages	41	0	0	0	1	11.61	8.96
Countdown Timers	31	2	1	0	0	9.10	6.42
Total	218	12	4	1	1	13.12	12.37

Table 4.1: The distribution of correct (CR), incorrect (IC), partially correct (PC), missed (MS), and unsure (US) classifications, and the mean classification timings (M) and their standard deviations (SD) in seconds, of dark patterns in our study.

We also compare the timings in regard to the *menu interfaces*.

Further, we compare the timings in regard to the different menu interfaces we employed during the study. The *Sidebar* was slightly faster than the *Bubbles* style (M=11.69, SD=13.22 vs. M=13.18, SD=8.11). Further, the *Low-To-High* task took the participants the longest in both styles (M=17.86, SD=10.72), although the *High-To-Low* style was not far off (M=13.3, SD=8.02). *Only-High* and *Only-Low* took the least classification time (M=7.24, SD=5.88 and M=8.25, SD=11.62). Interestingly, although our timing measurements suggest that the *Searchbar* was slightly quicker than *Low-To-High*, it was much slower than the comparable *Only-High* and *Only-Low* tasks, where users also had to only select one category.

We include all measurements in Table 4.2. The sidebar mean and standard deviation (M=11.29, SD=12.82) and the overall mean and standard deviation (M=13.12, SD=12.37) have to be considered to be skewed upwards, as the *Searchbar* was used by every participant, and therefore contributed the most values.

	Sidebar	Bubbles	
Only-High	M=4.24 SD=3.11	M=10.23, SD=8.64	M=7.24, SD=5.88
High-To-Low	M=11.32, SD=6.36	M=15.35, SD=9.69	M=13.3, SD=8.02
Only-Low	M=8.92, SD=17.96	M=7.59, SD=5.28	M=8.25, SD=11.62
Low-To-High	M=16.16, SD=13.62	M=19.57, SD=8.82	M=17.86, SD=10.72
Searchbar	M=15.81, SD=23.04	-	M=15.81, SD=23.04
	M=11.69, SD=13.22	M=13.18, SD=8.11	M=13.12, SD=12.37

Table 4.2: A comparison of the mean classification timings (M) and their standard deviations (SD) for our new menu interfaces, in seconds.

The 8 participants had mean classification timings between 6 and 22 seconds per dark pattern element. We found no correlation between their classification speeds and their age, the hours they spent online daily, their highest academic degree, or their perceived familiarity with the topic of dark patterns classification. However, as we only had a small number of participants, this may not be generalizable.

There was no correlation between classification time and age, hours spent online, academic degree, or familiarity with dark patterns classification.

The *Book of Dark Patterns* was opened a total of 24 times by all participants (M=2.86 times on average, SD=1.77), and closed after M=15.46 seconds on average, although some participants explored it much longer than others (SD=20.65). One participant didn't use the book at all, and therefore didn't answer any questions regarding it in the questionnaire. The *Tooltips* were hovered over and activated by every user at least once, although one user accessed them especially often (42 times), and the rest less so (M=9.75, SD=13.38).

We also measured the use of the *Book of Dark Patterns* and the *Tooltips*.

Questionnaire Results

In the following, we present the results from the 5-point Likert scale type questions of the questionnaires. All results can be found in Figure 4.3. The full questionnaires themselves can be found in Figure B.2. Because of our small size of participants, the results have to be taken into account with a lower generalizability. However, they still demonstrate interesting tendencies.

We start with a comparison of the four main selection tasks. The selection task *Only-High* was generally understandable (Q1, M=1.25, SD=1.09), similarly as much as *High-To-Low* (Q1, M=1.25, SD=0.82), and *Low-To-High* (Q1, M=1.25, SD=0.82). *Only-Low* was understandable even a bit more (Q1, M=1.75, SD=0.66). Participants found *Only-Low* to be the most intuitive (Q2, M=1.75, SD=0.43), and *Low-To-High* the least intuitive (Q2, M=0.88, SD=0,93). They further stated, that they enjoyed navigating *High-To-Low* the most (Q3, M=1.38, SD=0.70), and *Low-To-High* the least (Q3, M=0.88, SD=1.27).

The participants generally enjoyed the selection tasks, especially *Only-Low*, but less *Low-To-High*.

They found *Only-Low* to be the least frustrating (Q4, M=-1.63, SD=0.48), and, interestingly, *Only-High* the most (Q4, M=-1, SD=1.22), although all four selection tasks were generally not found frustrating.

The participants could most easily find and most quickly select the categories for *Only-Low*, and had most difficulties finding it for *High-To-Low* and selecting it for *Low-To-High*.

The participants could most easily find the category they were looking for in *Only-Low* (Q5, M=1.13, SD=0.93), and, notably, had most difficulties finding it in *High-To-Low* (Q5, M=0.13, SD=1.27). The interactive options of the UIs were generally regarded as helpful, although with little differences (Q6, M=0.88 to 1.13, SD=1.26 to 1.36). However, we observed that this question in context of selection tasks was generally hard to understand. Further, participants stated that *Only-High* and *Only-Low* had the most helpful designs for quick selections (Q7, M=1.5, SD=0.86 and M=1.5, SD=0.71), and *Low-To-High* the least helpful (Q7, M=0.88, SD=0.78). The perceived differences in quickly finding and selecting categories match our measured timings.

High-To-Low and *Low-To-High* were regarded as most promising for a better understanding of dark patterns classification.

Notably, the most participants could imagine, that playing with *High-To-Low* and *Low-To-High* may lead to a better understanding of the classification of dark patterns after longer periods of time (Q8, M=1.38, SD=0.99 and M=1.38, SD=0.49). In this regard, *Only-High* and *Only-Low* were rather seen as neutral (Q8, M=0.13, SD=1.54 and M=0.13, SD=0.93).

Only-Low was favored for novices, and *Low-To-High* for experts.

Further, the participants stated neutral opinions about *Only-High* being suitable for novices or experts (Q9, M=0.38, SD=1.32 and Q10, M=0.25, SD=0.83). For *Low-To-High*, they stated only slightly more positive opinions regarding expert use (Q9, M=0.38, SD=0.87 and Q10, M=0.63, SD=0.70). Notably, they stated that *Only-Low* is more suitable for novices than for experts (Q9, M=1.13, SD=1.05 and Q10, M=0.13, SD=0.93), and the reverse for *High-To-Low* (Q9, M=0.63, SD=0.99 and Q10, M=1, SD=0.71).

The *Searchbar* task was received with mixed opinions.

The *Searchbar* was a special kind of an *Only-Low* selection task, and graded by all participants at the end of the study. It was generally received with mixed opinions. It was understandable (Q1, M=1.5, SD=0.5), but less intuitive than any other task (Q2, M=0.75, SD=1.39). It also was by far the least enjoyable and most frustrating task (Q3, 0.25, 1.79

and Q4, 0.125, 1.27) in navigating its UI, although still neutrally graded. The question about easily finding the correct category from earlier was split into two for the *Searchbar*: The participants stated rather neutral opinions about still being able to easily determine and choose the category (Q5, M=0.5, SD=1.22), and finding it in the selection (Q6, M=0.88, SD=0.93), which were both lower rated than any other selection task in their corresponding question. Also, the *Searchbar* was rated to be the least well designed for enabling a quick category selection (Q7, M=0.88, SD=1.45). However, *Low-To-High* had the same, lowest mean result (0.88), with a smaller standard deviation (0.78). Our timing measurements support these opinions. Further, the *Searchbar* was seen as the least promising for a better understanding of dark pattern classification (Q8, M=-0.13, SD=1.17). Finally, participants stated it to be the least suitable for novices (Q9, M=-1.25, SD=0.83) and the most suitable for experts (Q10, M=1.25, SD=0.83) of all selection tasks.

The *Sidebar* was generally favored over the *Bubbles* UI style.

Next, we take a look at the results from the *Sidebar* and *Bubbles* interface styles. The *Sidebar* was rated as more intuitive to use than the *Bubbles* (Q1, M=1.13, SD=1.05 vs. M=0.25, SD=1.39), more enjoyable to navigate (Q2, M=0.5, SD=1.12 vs. M=0.13, SD=1.27), and less frustrating (Q3, M=-0.63, SD=1.41 vs. M=-0.38, SD=1.41). However, the *Sidebar* was rated as taking away more space unnecessarily than the *Bubbles* (Q4, M=-0.25, SD=1.39 vs. M=-0.88, SD=1.27), although both were rated negatively in this regard. Notably, the participants stated they could find the category a lot easier with the *Sidebar* than with the *Bubbles* (Q5, M=0.88, SD=1.17 vs. M=-0.25, SD=1.39).
Further, the participants found the interactive options offered by the *Sidebar* to be less helpful than the ones by the *Bubbles* (Q6, M=0.63, SD=1.32 vs. M=1.88, SD=0.33). However, this doesn't necessarily show an advantage of the latter, as the *Bubbles* only worked by making the category names appear when hovering over the buttons, which the *Sidebar* didn't need to do. Interestingly, the participants also found the selection to be much quicker with the *Sidebar* than with the *Bubbles* (Q7, M=0.75, SD=1.39 vs. M=-0.25, SD=1.48). The perceived differences in quickly finding and selecting categories match our measured timings.

Also, the participants found the *Sidebar* to be more promising for a better understanding of dark patterns classification (Q8, M=1.25, SD=1.3 vs. M=0.88, SD=0.78). Finally, the participants found the it to be more suitable for both novices and experts than the *Bubbles* (Q9, M=1.25, SD=0.81 vs. M=-0.5, SD=1.32 and Q10, M=1, SD=0.87 vs. M=0.38, SD=0.7). In this regard, the *Bubbles* were seen as more suitable for experts.

The *Book* and the *Tooltips* were similarly liked.

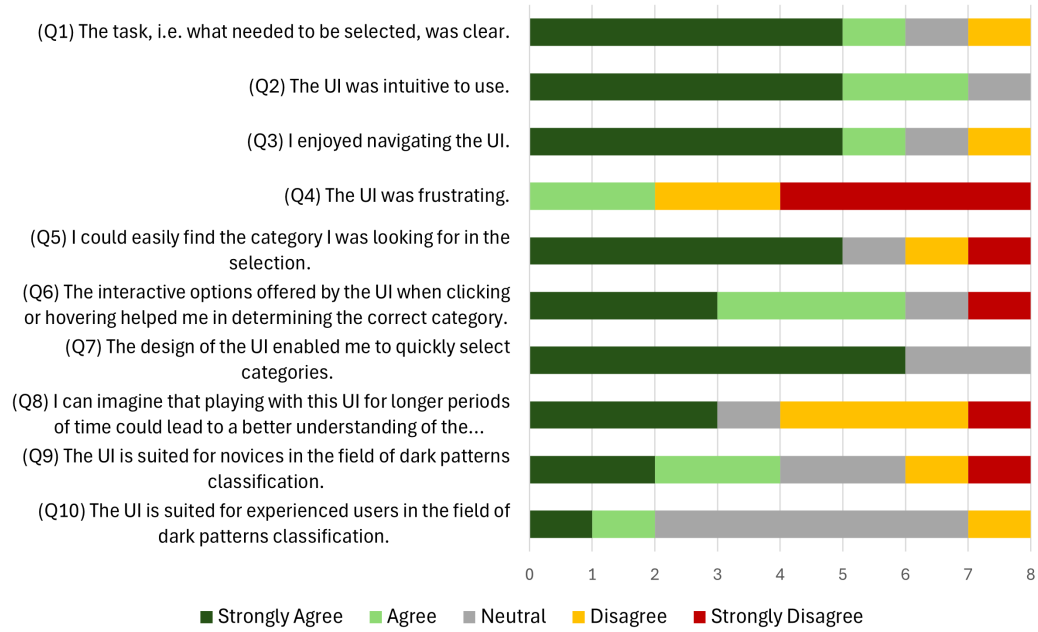
Next, we examine the results from the *Tooltips* and the *Book of dark patterns*. The *Tooltips* were graded as slightly more intuitive than the *Book* (Q1, M=1.13, SD=1.05 vs. M=1, SD=0.76), slightly less enjoyable in their navigation (Q2, M=0.38, SD=1.22 vs. M=0.58, SD=1.05), but also slightly less frustrating (Q3, M=-0.75, SD=0.83 vs. M=-0.58, SD=1.29). Surprisingly, participants were able to find the *Tooltips* less easily than the *Book* (Q4, M=1.13, SD=1.05 vs. M=1.71, SD=0.45). Also, they stated that both equally didn't take unnecessarily much space on the page that could have been used otherwise (Q5, M=-1, SD=1.22 vs. M=-1, SD=0.93), albeit the two being implemented in different locations.

Although the *Tooltips* offered a quicker way of finding information, the *Book* was preferred for its ease of use in determining the correct category.

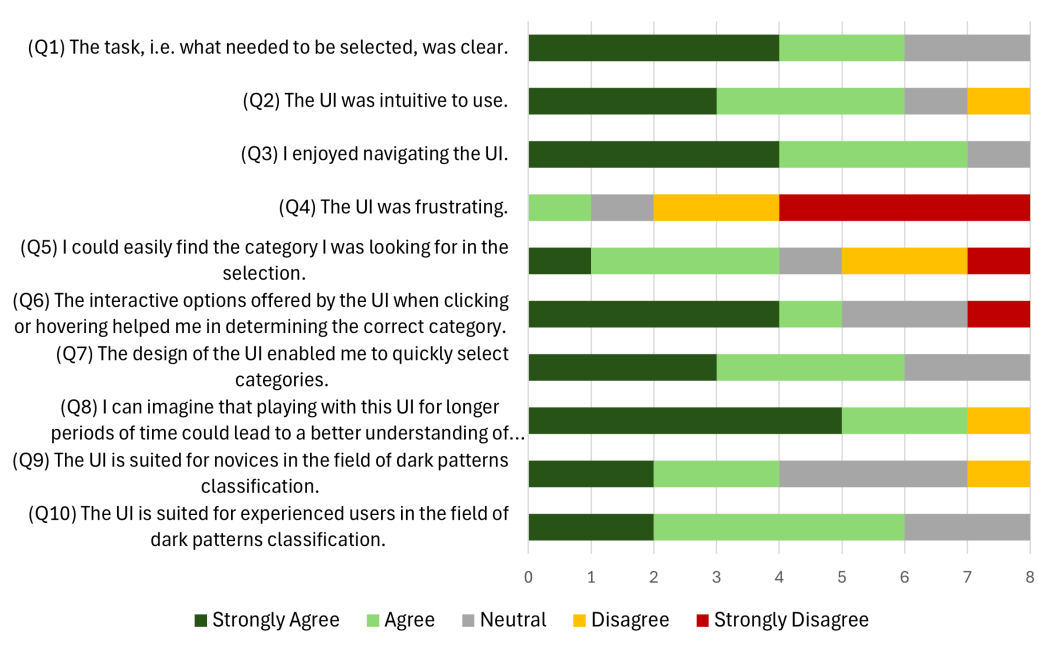
Notably, the participants stated, that they found the information they were looking for with more difficulty with the *Tooltips*, than with the *Book* (Q6, M=0.5, SD=1.32 vs. M=1.29, SD=0.7). However, they also stated that the design of the *Tooltips* helped more in getting to the desired information quickly, than the *Book* (Q7, M=0.5, SD=1.32 vs. M=0.14, SD=1.12). Further, the *Tooltips* were stated as being less helpful in determining the correct category than the *Book* (Q8, M=0.88, SD=1.27 vs. M=1.43, SD=0.73), and less promising for a better understanding of dark pattern classification (Q9, M=0.88, SD=1.45 vs. M=1.57, SD=0.73).

Both were regarded as more suited for novices.

Finally, the *Tooltips* were seen as less suited for novices than the *Book* (Q10, M=0.88, SD=1.27 vs. M=1.29, SD=0.7), but also far less suited for experts (Q11, M=0, SD=1.12 vs. M=0.71, SD=0.88). In general, both were regarded as less suited for experts than for novices, which makes sense, as they offer hints for the correct solution, which experts may need less.

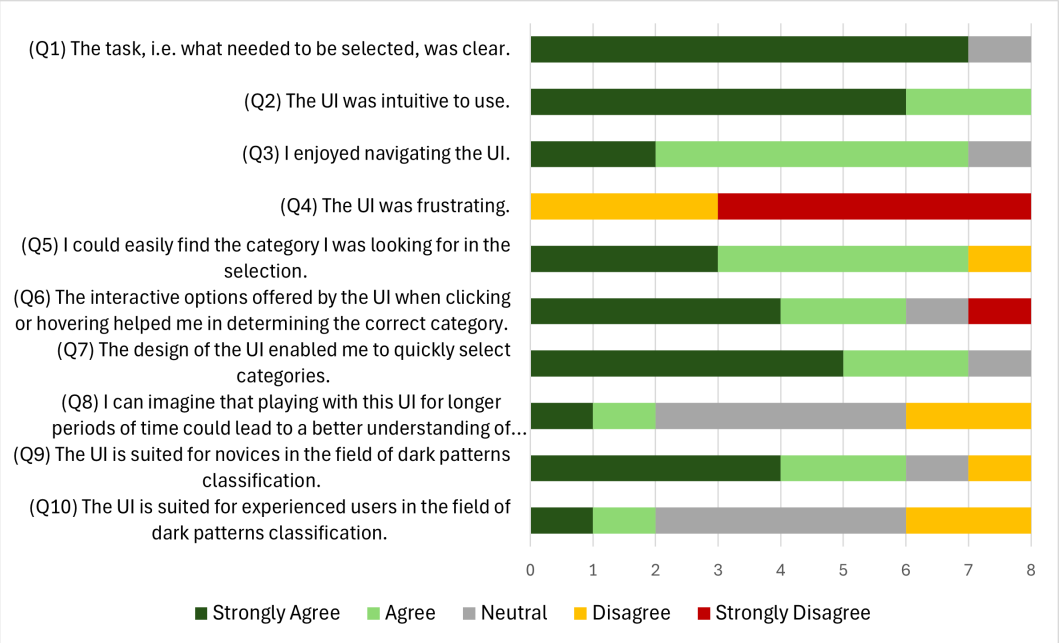


(a) Only-High

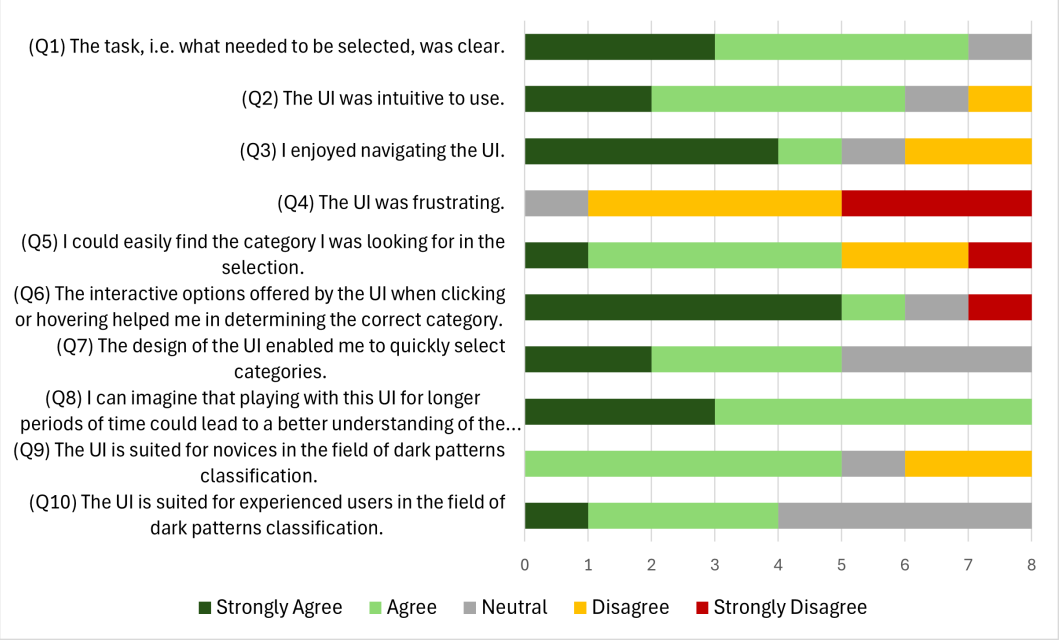


(b) High-To-Low

Figure 4.3: The results from the 5-point Likert scale type questions from the questionnaires.

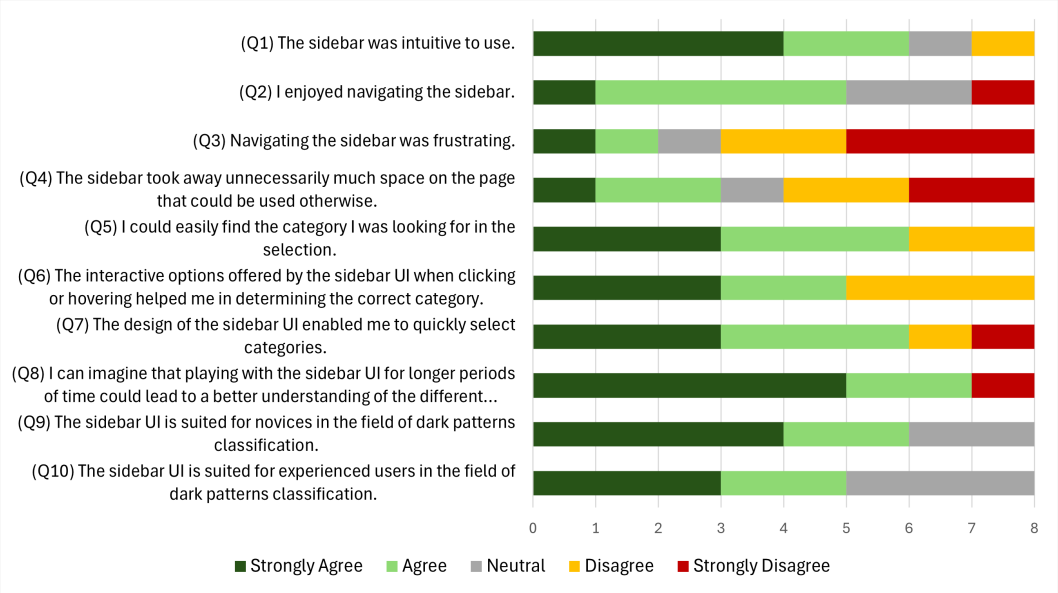


(c) Only-Low

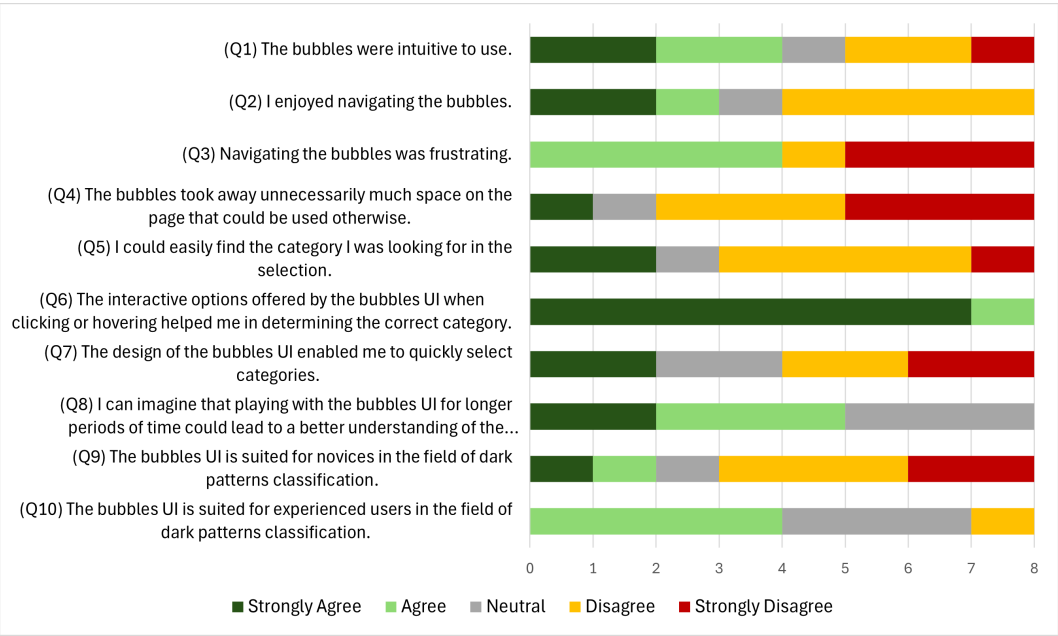


(d) Low-To-High

Figure 4.3: The results from the 5-point Likert scale type questions from the questionnaires. (cont.)

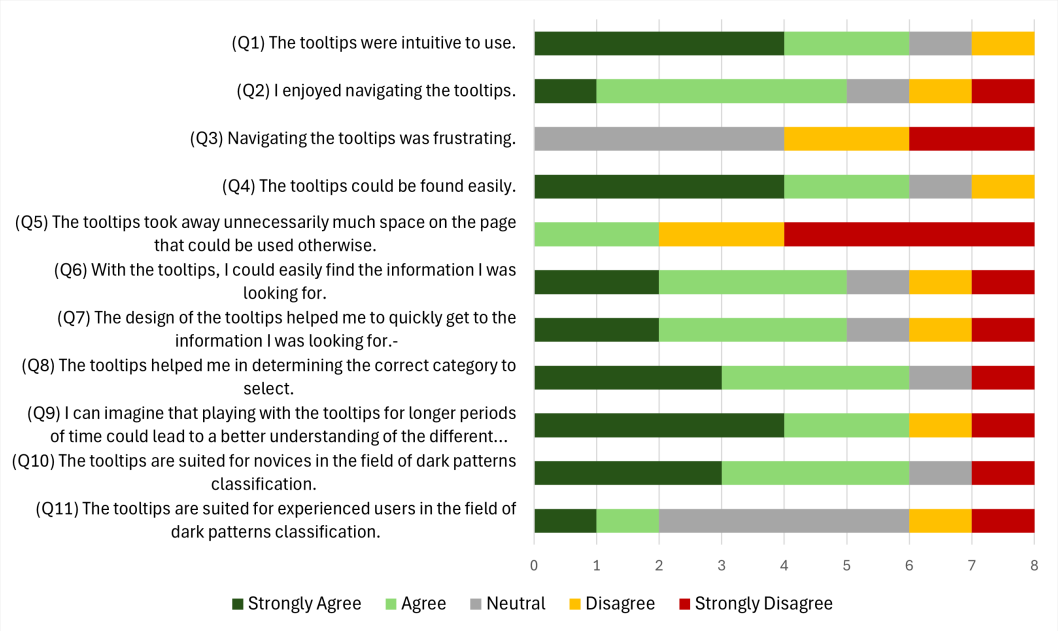


(e) Sidebar

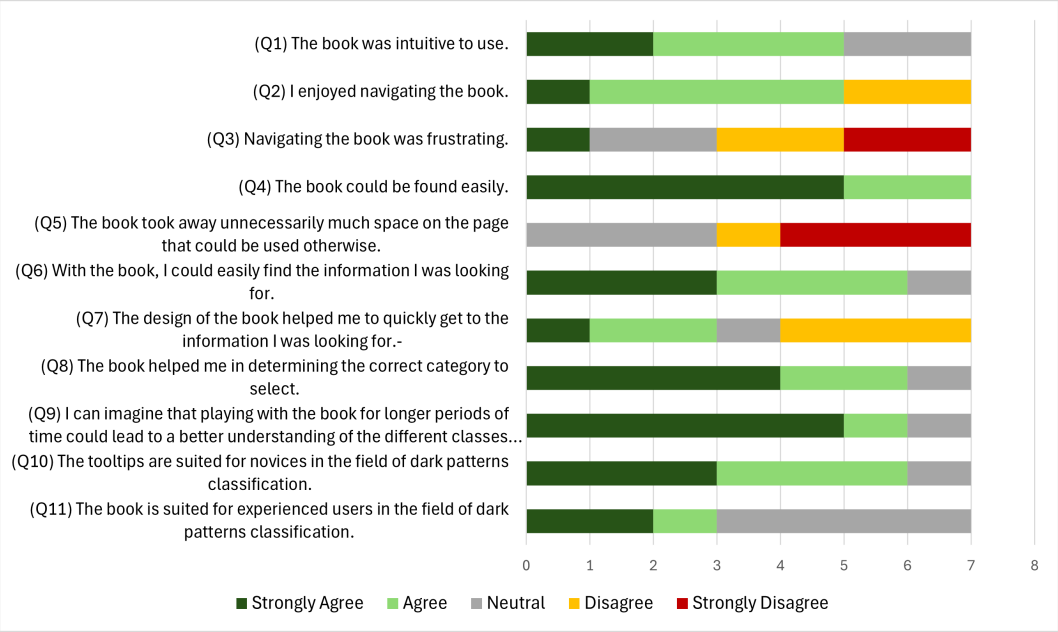


(f) Bubbles

Figure 4.3: The results from the 5-point Likert scale type questions from the questionnaires. (cont.)

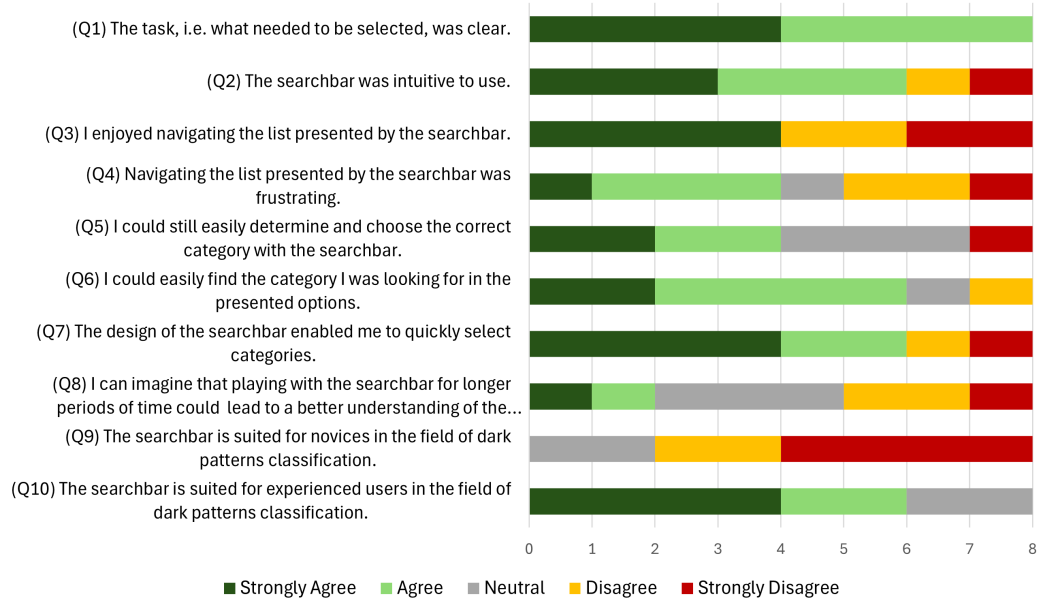


(g) Tooltips



(h) Book of Dark Patterns

Figure 4.3: The results from the 5-point Likert scale type questions from the questionnaires. (cont.)



(i) Searchbar

Figure 4.3: The results from the 5-point Likert scale type questions from the questionnaires. (cont.)

Overall, the participants liked the selection tasks employing the ontology hierarchy the most.

At the end of the questionnaire, the participants stated, how they would rank the 5 selection tasks, regarding three aspects. We examine the average of these rankings: Regarding *Ease of Use*, *Only-High* was ranked best (M=4.38, SD=1.11), followed by *Only-Low* (M=3.38, SD=1.22), then *High-To-Low* and *Low-To-High* (M=2.75, SD=0.97 and M=2.75, SD=1.2), and *Searchbar* worst (M=1.75, SD=1.09). Regarding the *Difficulty* of the task, *Searchbar* was ranked most difficult (M=4.18, SD=1.54), followed by *High-To-Low* and *Low-To-High* (M=3.38, SD=0.7 and M=3.38, SD=0.99), then *Only-High* (M=2.25, SD=1.48), and *Only-Low* the easiest (M=1.88, SD=0.78). *Overall*, *High-To-Low* was ranked best (M=4.38, SD=0.86), followed by *Low-To-High* (M=3.13, SD=1.17), then *Only-High* and *Only-Low* (M=2.75, SD=1.39 and M=2.75, SD=1.09), and *Searchbar* worst (M=2, SD=1.32).

The favorite combinations show a preference for hierarchy employing tasks, and a balanced one regarding styles and definition representations.

Finally, the participants stated their favorite combinations of selection tasks, interface styles, and category definition representations. *High-To-Low* + *Bubbles* + *Book* was selected three times (37.5%), *High-To-Low* + *Sidebar* + *Tooltips* two times (25%), *Low-To-High* + *Sidebar* + *Tooltips* also two times (25%), and *High-To-Low* + *Sidebar* + *Book* once (12.5%). All participants chose selection tasks employing the ontology hierarchy, which matches their "Overall" rankings. 5 of 8 participants (62.5%) chose the *Sidebar* over the *Bubbles*. Half chose the *Tooltips*, and half the *Book*. Interestingly, the *Sidebar* was paired with the *Tooltips* in 80% of the cases, and only once with the *Book*.

4.2.3 Qualitative Analysis

In the following, we present the participants' compiled answers for the free-text input fields in our questionnaires, their comments during the study, and further notes we made during it. In our data, the participants were labeled with random user id's to ensure anonymity. The numbers we use to address participants in this section are arbitrary and correspond to these random id's.

The participants enjoyed the study.

When we asked participants how they found the study, they stated that they generally had enjoyed it. They were interested in seeing the further development of the learning game by Fiedler [2024].

Conveying the ontology hierarchy was important to the participants.

The selection task *Only-High* was often stated to be too easy and not helpful for teaching dark pattern categories (*"I didn't like it, as it didn't include the hierarchy. Also, the high-level categories are too broad." - Participant 2*). Further, although *Only-Low* was liked for employing lower level categories, it was similarly criticized for its lack of the ontology hierarchy:

"It does not force the user to classify the dark patterns into a high level category, which means that the user is not building up a hierarchy or classification in their head while playing the game."

—Participant 7

High-To-Low was largely favored by the participants. However, some struggled to find the correct lowest level category in its selection. They tended to go back and forth in the hierarchy to find the correct category. Mostly, they didn't use the book for this, which would have offered a similar navigation. One participant even found the task tedious:

Still, the selection tasks employing the hierarchy were not universally liked.

"Sometimes it's a bit odd to figure out which high level category a specific low level dark pattern category sits in."

—Participant 5

Although *Low-To-High* was often regarded as too complicated, participants stated they liked it including the hierarchy, while still reducing choice overload:

"I think starting at the low level directly acquaints the user for a specific kind of dark pattern reducing choice overload."

—Participant 3

The *Searchbar* was regarded as an expert menu interface, as the user had to already know the name of the correct category (*"If the user already knows what to select, and hence what to type in the search bar to reduce the list of choices presented, it does make the task easier." - Participant 4*). Some participants forgot certain category names or remembered them falsely, having to scroll and search through the whole list of lowest level dark patterns. As an addition, it was suggested by two participants, that input text could also be matched

Some participants found the *Searchbar* difficult to use, as they didn't remember all category names.

to higher level categories, displaying them in the High-To-Low style, and then being able to get to the lowest-level categories over the hierarchical selection.

The *Bubbles* were often disliked for their lack of visible category names.

The *Bubbles* UI style was equally liked and disliked by the participants. Most complains regarding it were, that the buttons didn't show the category names directly, having to hover over most options until finding the correct one:

> *"I did not enjoy the bubbles. Hovering over the buttons takes too long."*
>
> *—Participant 8*

However, one participant stated the assumption, that after some time, a learning effect may set in, as the user mentally connects the emojis with the categories.

Although the *Bubbles* could obscure elements, this was not received negatively.

We noticed, that when categorizing elements with the *Bubbles* UI, participant 6 clicked on an other element multiple times to make the interface disappear, so they could see the element better. However, they didn't note this negatively for the *Bubbles* style - some participants actually liked it for its placing on the page (*"I think the strength is that the hierarchical pie menu allows the user to think about the dark pattern in the context it is present. It does not distract and take your focus away from the dark pattern at hand." - Participant 4*). Some even asked for the buttons of the bubbles to be larger.

Further, two participants stated an idea for a multi-layer radial menu, that had strong similarities with one of our paper prototypes (see Figure A.3), as an improvement of the *Bubbles* (*"[...] probably a multi-level pie menu that can be navigated similar to how it is in Blender or other 3D modelling software." - Participant 6*).

The *Bubbles'* greatest weakness was the *Sidebar's* greatest strength.

In comparison to the *Bubbles*, for the *Sidebar*, the participants stated, that them being able to read the category names without hovering over the buttons was its main strength (*"One strength is that you can read and find categories by their names. I found waiting for the names with the bubbles interface annoying." - Participant 3*).

Notably, one participant would have liked an accordion type interface more than a sequential one for the *Sidebar*, similarly described to the one we designed in Figure A.1.

The *Sidebar* and *Tooltips* combined took away much space.

Further, some stated that the *Sidebar* takes too much space on the page (*"It takes way too much space on the side" - Participant 3*), especially when combined with the *Tooltips* (*"I would try to separate them [the tooltips] from the selection UI." - Participant 4*), as then, the list of options in the *Sidebar* had less space. However, the tooltips were generally liked for their quick accessibility (*"I liked that I could quickly access the tooltips" - Participant 2*).

The *Book* was generally liked for its hierarchical structure and allocated place for information.

> *"I love the book! I think it's better for understanding dark patterns, as it has more space to explain them, and it also features the hierarchy connections, which may be useful. The book is so cute!"*
>
> *—Participant 8*

The *Book* was often used to look up definitions or hierarchy relations.

The *Book* was often opened to look up a category definition. For example, participants 4 and 6 used the book to find the definition for *Trick Questions*, read it, and then categorized it correctly. Further, the *Book* was used multiple times to find higher level categories for lowest level ones in the *Low-To-High* task.

We also identified, that two participants used the arrow buttons at the sides of the *Book* incorrectly, as they assumed they were meant to go back and forth in the hierarchy, and not sideways. However, they did not criticize this openly.

The participants had further ideas for improvements of the *Book*.

Some participants stated, that the *Book* could also use a searchbar for category names to quickly look up category definitions. Two participants further proposed the idea of "bookmarks" for the *Book*, that was very similar to the one from our prototype (see Figure 3.5).

Some participants shared interesting ideas for new menu interfaces.

Interestingly, two participants had the idea of combining the *Bubbles* interface with the *Searchbar*, to enable both novices and expert users to make quick selections (*"Maybe the search bar and the bubbles could be combined, so that novices, as well as experts, can choose categories quickly, and at the same time." - Participant 1*).

The emojis and color scheme were received well.

Some participants enjoyed the inclusion of emojis (*"The emojis are great. It must have been great fun to pick one for every category." - Participant 2*). Indeed, we enjoyed choosing the emojis. The participants also stated they liked the color scheme for the categories from different hierarchy levels.

We observed an anticipated learning effect.

Most participants stated, that after a few levels, they had figured out the categories for the elements and were only searching for the correct options in the selections, not thinking about which are correct anymore. However, this was anticipated, and we aimed to circumvent this by having different orders of selection tasks, styles and definition representations. We also observed that this learning effect helped most participants in the *Searchbar* task, which otherwise would have been even more difficult to use.

4.2.4 Discussion & Interpretation

In this section, we discuss and interpret the study results to answer our research questions.

RQ-1: Which menu interfaces are most preferred?

High-To-Low and *Low-To-High* are the preferred selection tasks.

Some of the quantitative results, like slower perceived and actual selection times, or less stated enjoyment, suggest, that participants had problems with the selection tasks employing the ontology hierarchy (*High-To-Low* and *Low-To-High*). However, the participants still exclusively picked these two for their favorite combinations, and also placed them as the two best tasks in the *Overall* ranking. This could be, because the participants had the goal of the study in mind, aiming to find menu interfaces suitable for the learning of dark pattern classification. In this regard, they stated

High-To-Low and *Low-To-High* to be most promising for a better understanding of dark patterns classification. Despite ranking them most difficult in *Difficulty* and worst in *Ease of Use* (before *Searchbar*), they showed a great interest in conveying the ontology hierarchy with the selection task.

Further, some participants expressed a liking in directly selecting lowest level categories, instead of starting at the high-level ones. Still, it is not clear, if *High-To-Low* or *Low-To-High* is the most preferred selection task. This is surprising to us, as at the start of our research, we presumed that the *High-To-Low* task may have a major drawback in starting at the highest, most abstract hierarchy level. However, the *Searchbar* task stood out as being the most disliked, mostly for its difficulty. It was often regarded as being better suited for additional interfaces for experts, but not so much for an educational purpose.

Both the *Bubbles* and the *Sidebar* have respective advantages.

Again, some of the quantitative results, like slower perceived and actual selection times, or less stated enjoyment, suggest, that the *Sidebar* was preferred over the *Bubbles*. However, the participants also stated some aspects they liked about the *Bubbles*, that the *Sidebar* didn't employ, like taking away less space, or the placing of the interface. Also, both were chosen almost equally often in the favorite combinations. It is therefore not clear, if the *Bubbles* or the *Sidebar* are the preferred interface style.

Similarly, both the *Book* and the *Tooltips* have respective advantages.

The *Book of Dark Patterns* and the *Tooltips* were similarly liked. Both were chosen exactly equally often in the favorite combinations. Further, mixed quantitative and qualitative results suggest that both have respective advantages that make them useful: most notably, a better accessibility for the *Tooltips* (directly inside the selection interface), and a richer information representation for the *Book*, including hierarchical relations. It is therefore not clear, which one is preferred. However, they don't have to be exclusively used, as both employ quite different methods of conveying information.

We suggest to mix the preferred menu interfaces in the game.

Furthermore, as both the *High-To-Low* and the *Low-To-High* selection tasks, both the *Bubbles* and the *Sidebar* styles, and both the *Tooltips* and the *Book of Dark Patterns* have their respective advantages and disadvantages, it may be a valuable approach to include all of them, or alternate between them, in the regular game. For example, we included a button in our implementation to seamlessly switch between the interface styles. The selection tasks could also be alternated between levels, making the gameplay more varied - the implementation for such a function is also already built in. Further, when using the *Sidebar*, both category definition representations could be allowed at the same time, albeit with the caveat of possibly making the game too easy.

RQ-2: Which menu interfaces allow for the quickest selection?

Surprisingly, *Linear* style interfaces were quicker than *Bubbles* style ones.

Our timing measurements suggest, that the selection tasks with only one selection level (*Only-High* and *Only-Low*) are quicker to accomplish than the ones with multiple hierarchy levels (*High-To-Low* and *Low-To-High*). This is expected, as the latter require more selections. However, we were surprised to see that the *Sidebar* UI had quicker selection timings than the *Bubbles* UI, and participants perceived both the "finding" and the "selecting" of options slower with the latter. This is unexpected, as literature suggests that after finding an option, it is usually selected more quickly with radial menus than linear ones [Callahan et al., 1988; Samp and Decker, 2010]. Also, Samp and Decker [2010] assume that using radial layouts for menus with changing content order may worsen performance. However, we cannot replicate this with our timing measurements, either: The *Only-Low + Bubbles* menu interface, which involves subsets with a random order of options, has the second smallest average selection time.

RQ-3: Do players better understand specific dark patterns after classifying them with the new menu interfaces?

The participants likely reached a higher cognitive process dimension.

We are confident, that our participants not merely *remembered* the dark patterns, but also *understood* their workings in our study. For example, the participants looking up definitions in the *Book* or the *Tooltips* and then **correctly** guessing between two visually similar, but conceptually different categories like *Trick Questions* and *Confirmshaming*, is a strong indicator for this. Also, as noted in our discussion of RQ-1, the participants largely state the inclusion of lowest level categories (and the respective ontology hierarchies) to be better suited for conveying dark pattern classification, supporting our claim. Further, our participants guessed the categories correctly for 97% of all dark pattern elements, compared to 77.67% of correct guesses in the original study by Fiedler [2024] - although this has to be viewed with the caveat, that our study focused on the comparison of menu interfaces, therefore employing more obviously recognizable elements. Generally, we believe, that the participants reached a higher cognitive process dimension (from merely *recognizing* to *classifying*) in the conceptual knowledge dimension in Blooms Revised Taxonomy [Krathwohl, 2002] after playing through our study.

Another study may be beneficial for statements about a generally better learning effect.

However, the question remains, if a higher cognitive process dimension for the classifying game task also leads to an overall better recognition of and confidence dealing with dark patterns. This study alone may be not sufficient to answer, if a generally better learning effect can be inferred. A second user study might be beneficial to answer this, for example measuring data in the full implementation of the game (including the finding task), and employing tests for knowledge about dark pattern classifications.

Chapter 5

Summary and Future Work

In this chapter, we summarize our findings, note their limitations, and discuss potential directions for future research and improvements for the game.

5.1 Summary & Contributions

In this thesis, we contribute research on various menu interfaces for the classifying task of the learning game by Fiedler [2024].

We first explored different menu interfaces for the classifying game task.

Firstly, we explored different interface designs by prototyping on paper. Then, we derived design considerations from notable differences in these prototypes, namely selection tasks, interface styles and category definition representations. We implemented these menu interfaces into the existing game, together with a sophisticated data structure conveying the dark pattern ontology and its relations.

We evaluated our menu interfaces and found promising results.

Lastly, we evaluated our menu interfaces in a user study. We thoroughly examined our results on the various advantages and disadvantages of these interfaces. Notably, we find a clear preference for selection tasks employing the ontology hierarchy, and are confident that classifying dark patterns with our new menu interfaces leads to a better understanding of the workings and harms of specific dark patterns.

5.2 Limitations & Future Work

Our user study only had 8 participants, who all were familiar with dark patterns.

As mentioned, a major limitation was that we only included 8 participants in our user study, due to the limited scope of this thesis. This makes our results less generalizable. Also, all participants were familiar with the domain of dark patterns, as the focus of our study lay on the comparison of the menu interfaces.

A second user study may be beneficial for statements about a generally better learning effect.

A second user study with more participants unfamiliar with the subject, including tests for knowledge about dark pattern classifications and measuring data in the full implementation of the learning game, may be beneficial to further answer, if understanding dark pattern categories actually leads to an overall better recognition of and confidence regarding dark patterns.

It may be interesting to test the retained knowledge in long-term studies.

Further it may be interesting to test the participants retained knowledge on dark pattern categories after longer periods of time, to find if they still understand the workings of the dark patterns from the learned categories, as Fiedler [2024] similarly suggested.

The *Searchbar* may be explored further.

Notably, the *Searchbar* was received rather negatively. It may be interesting to explore it further in other contexts. For example, combining it with other menu interfaces, or expanding it to feature higher level categories (and related tasks) as well, as suggested by participants.

The *Book of Dark Patterns* may be expanded.

Moreover, the *Book of Dark Patterns*, that currently serves as a complete encyclopedia on dark pattern categories, could be further expanded to include more features in the game. For example, early on in our research on the prototypes, we had the thought of making the pages of the book "unlockable". In this version, the book would be empty at the start. When the user would play the regular game and find a dark pattern of a specific category for the first time, it would then reveal the categories page, and add a screenshot of the found pattern as the example picture. This may create a bonding of the category definition to the personal experience of discovering the corresponding dark pattern. The user would have another incentive to play the game, gradually filling out the book by finding more novel dark pattern types.

The ontology included in the game allows for new, sophisticated features.

Furthermore, the inclusion of the whole ontology of dark patterns in the game allows for further, new features. For example, the game could track the where the user has the least success in classifying specific dark pattern categories, and then recommend levels (or even generate levels) which majorly employ dark pattern elements from these categories, further enhancing learning success at the places the user has the most needs for it.

Appendix A

Paper Prototypes

In the following, we include all paper prototypes that were created in addition to the ones presented in Chapter 3.1, "Prototyping" at the beginning of our research.

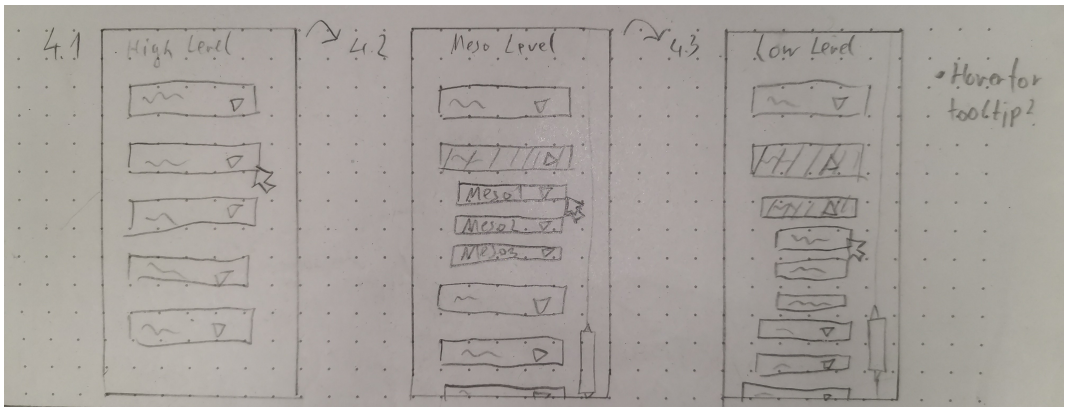


Figure A.1: A paper prototype in the style of an accordion[1] type menu interface. This can be regarded as a type of an *expandable linear menu* (see Chapter 2.3, "Viable Designs for Menu Interfaces").

[1] https://www.nngroup.com/articles/accordions-on-desktop/[Accessed:Nov.18,2024]

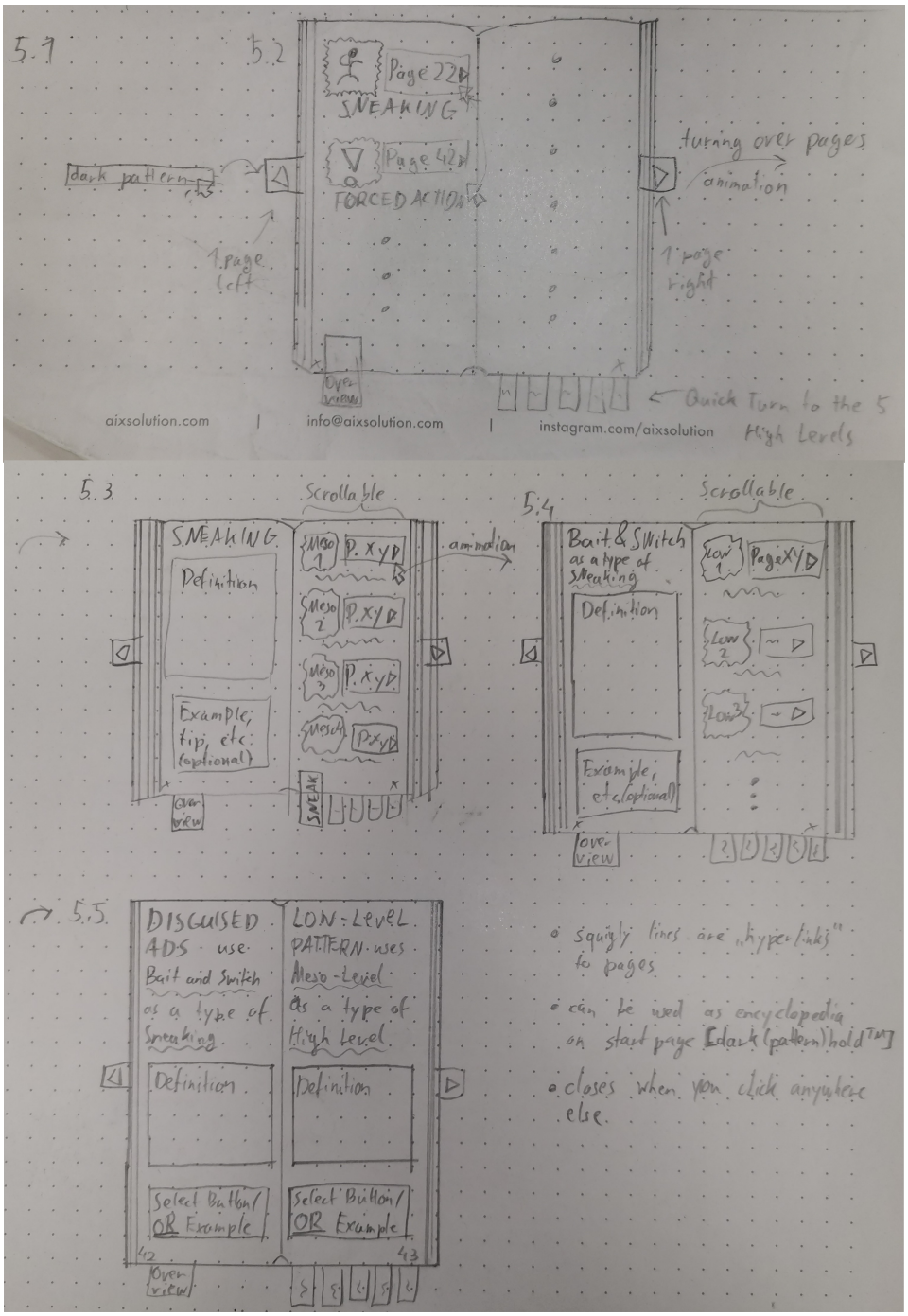


Figure A.2: A paper prototype of the *Book of Dark Patterns*. This started as an idea for another menu interface, but was reworked to serve as an alternative to the *Tooltips* definition representation (see Chapter 3.2.2, "Category Definition Representations").

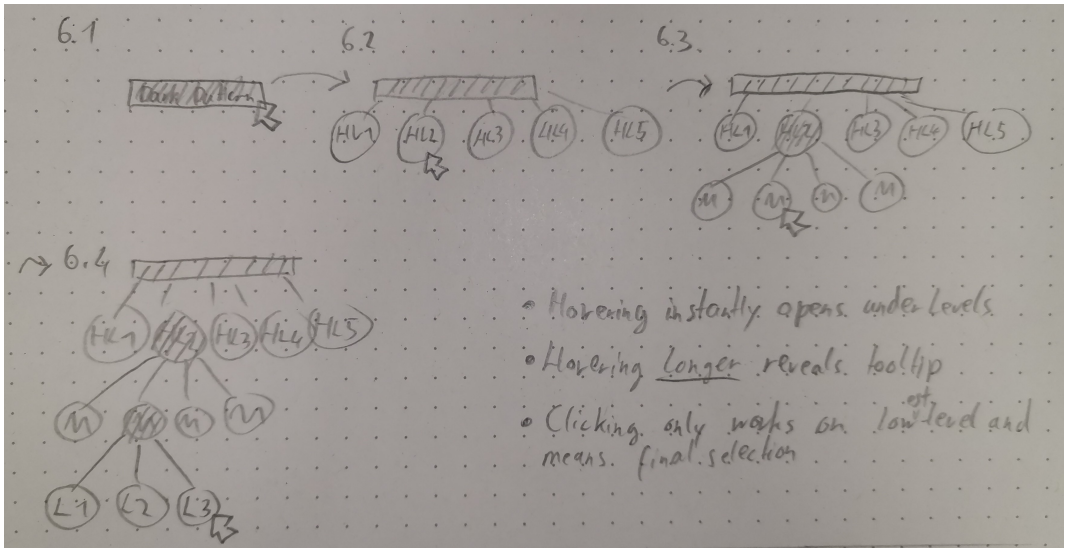


Figure A.3: A paper prototype in the style of an *extendable radial* type menu interface (see Chapter 2.3, "Viable Designs for Menu Interfaces"). The main idea here is to have previous hierarchy levels of the menu remain visible, and selecting lower hierarchy levels by hovering over options. As the levels expend in one direction, this doesn't look like a radial menu, but it implements the paradigms of one. A mouse click on an options means its final selection.

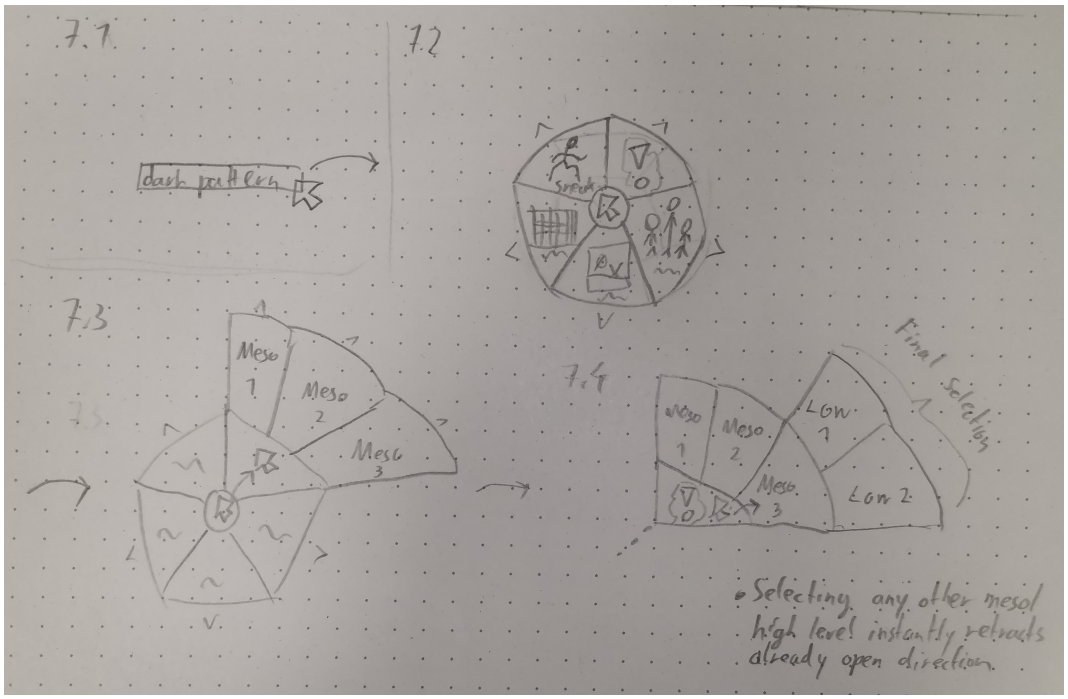


Figure A.4: Another paper prototype in the style of an *extendable radial* type menu interface (see Chapter 2.3, "Viable Designs for Menu Interfaces"). The main idea here is to have previous hierarchy levels of the menu remain visible, and having the lower levels extend in all directions like in [Samp and Decker, 2010]. If the options have to be clicked or only hovered over is not specified.

Appendix B

Menu Interfaces Study

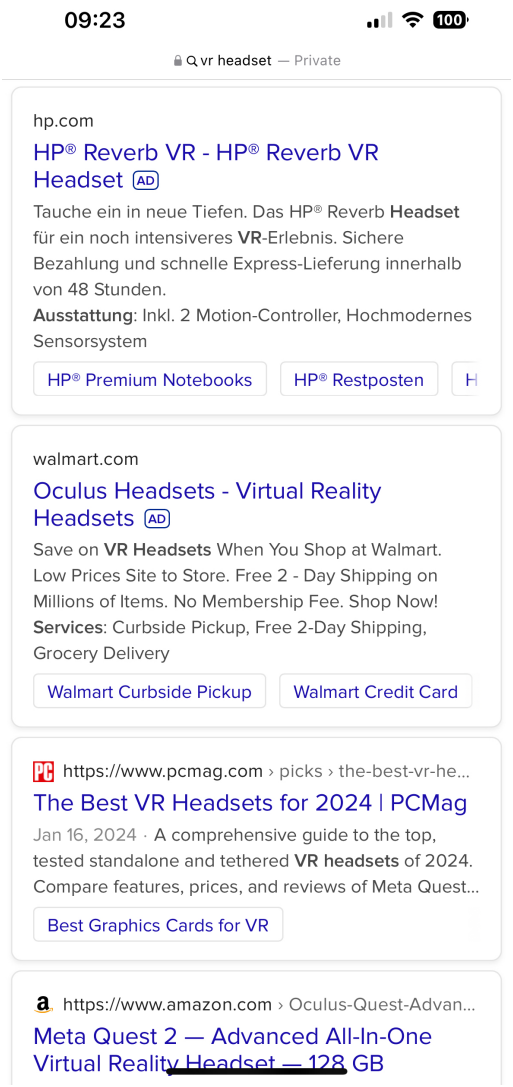
In the following, we present material used in the user study.

B.1 Image Assets

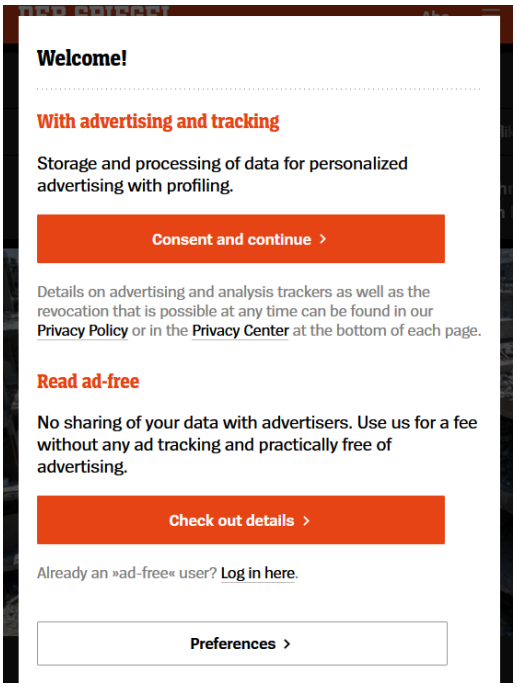
In the tutorial section of our user study, we used images from websites and from the original tutorial by Fiedler [2024] as examples for the defined categories. They are included in Figure B.1.

B.2 Questionnaires

The user study involved three questionnaires to evaluate the effectiveness of the menu interfaces and category definition representations. They are included in Figure B.2.

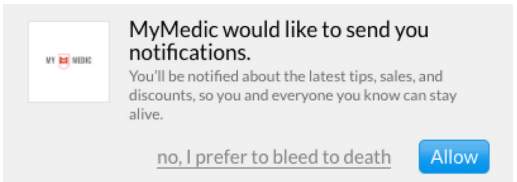


(a) An example for *Disguised Ads* from the original tutorial by Fiedler [2024]



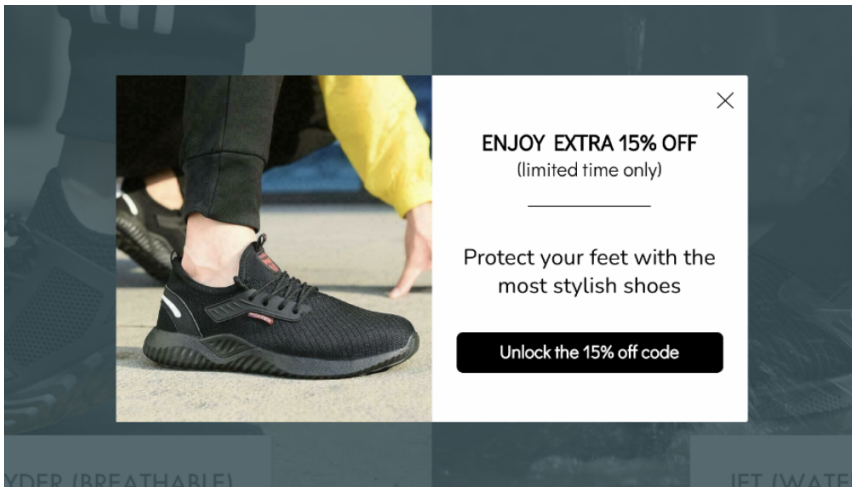
(b) An example for *False Hierarchy* from https://www.spiegel.de/ [Accessed on: Nov. 21, 2024]

(c) An example for *Trick Questions* from https://app.uxcel.com/lessons/dark-patterns-024 [Accessed on: Nov. 21, 2024]

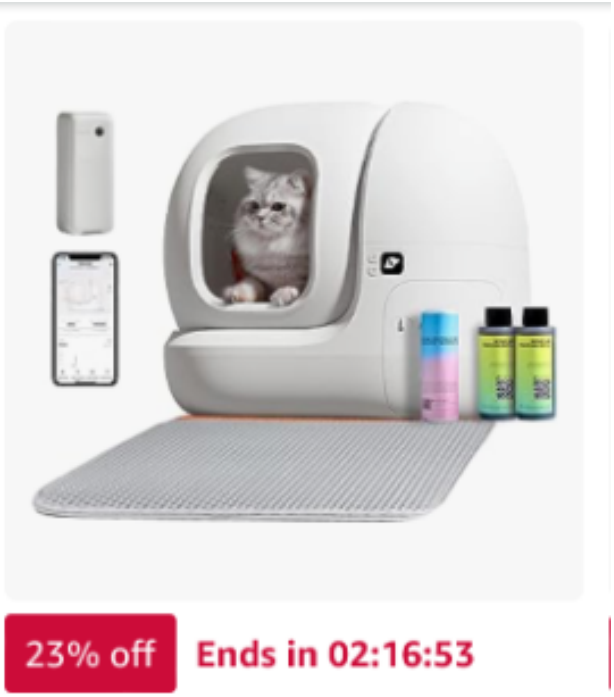


(d) An example for *Confirmshaming* from https://www.deceptive.design/types [Accessed on: Nov. 21, 2024]

Figure B.1: The images used in the tutorial section of our user study

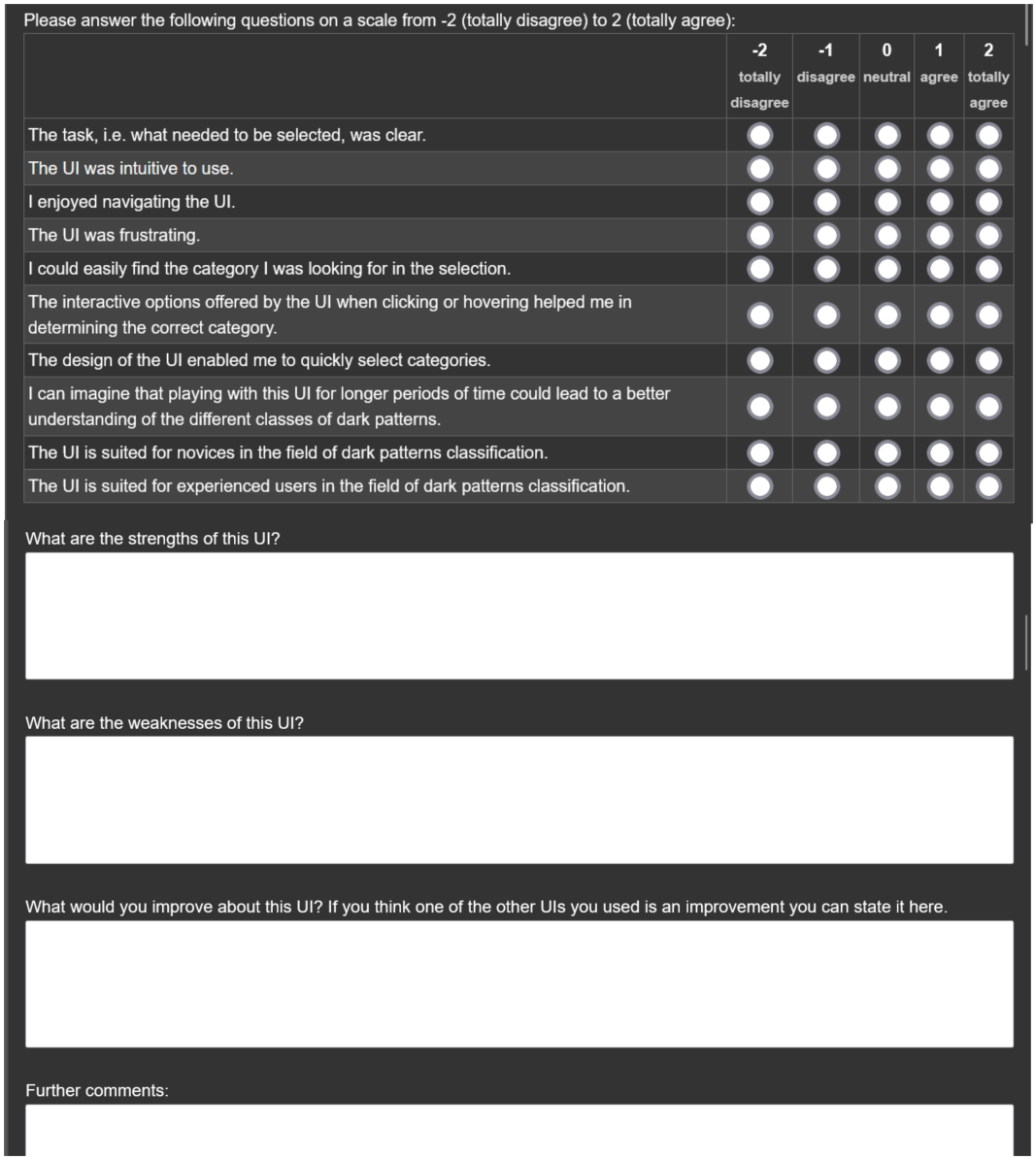


(e) An example for *Limited Time Messages* from https://www.optimonk.com/limited-time-offers/ [Accessed on: Nov. 21, 2024]



(f) An example for *Countdown Timers* from the original tutorial by Fiedler [2024]

Figure B.1: The images used in the tutorial section of our user study (cont.)

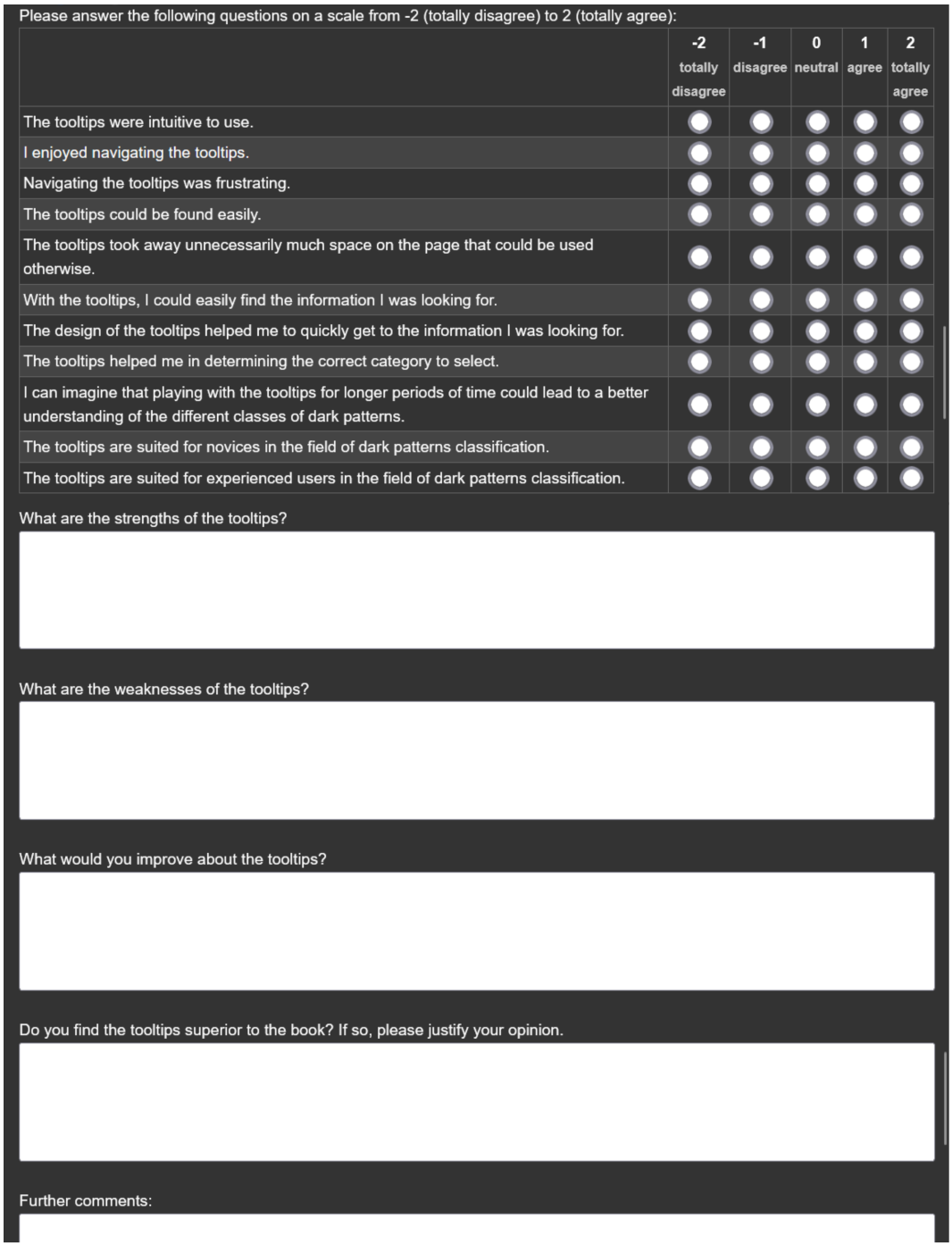


(a) The first questionnaire asked for the participants' opinions about the first 4 user interfaces they interacted with, focussing on the comparison of their selection tasks. The questions in the figure were asked 4 times in total, once for each UI the participant encountered, along with a picture (from Chapter 3.2.1, "Selection Tasks & Interface Styles") of the corresponding interface as a reminder.

Figure B.2: The three questionnaires included in our user study

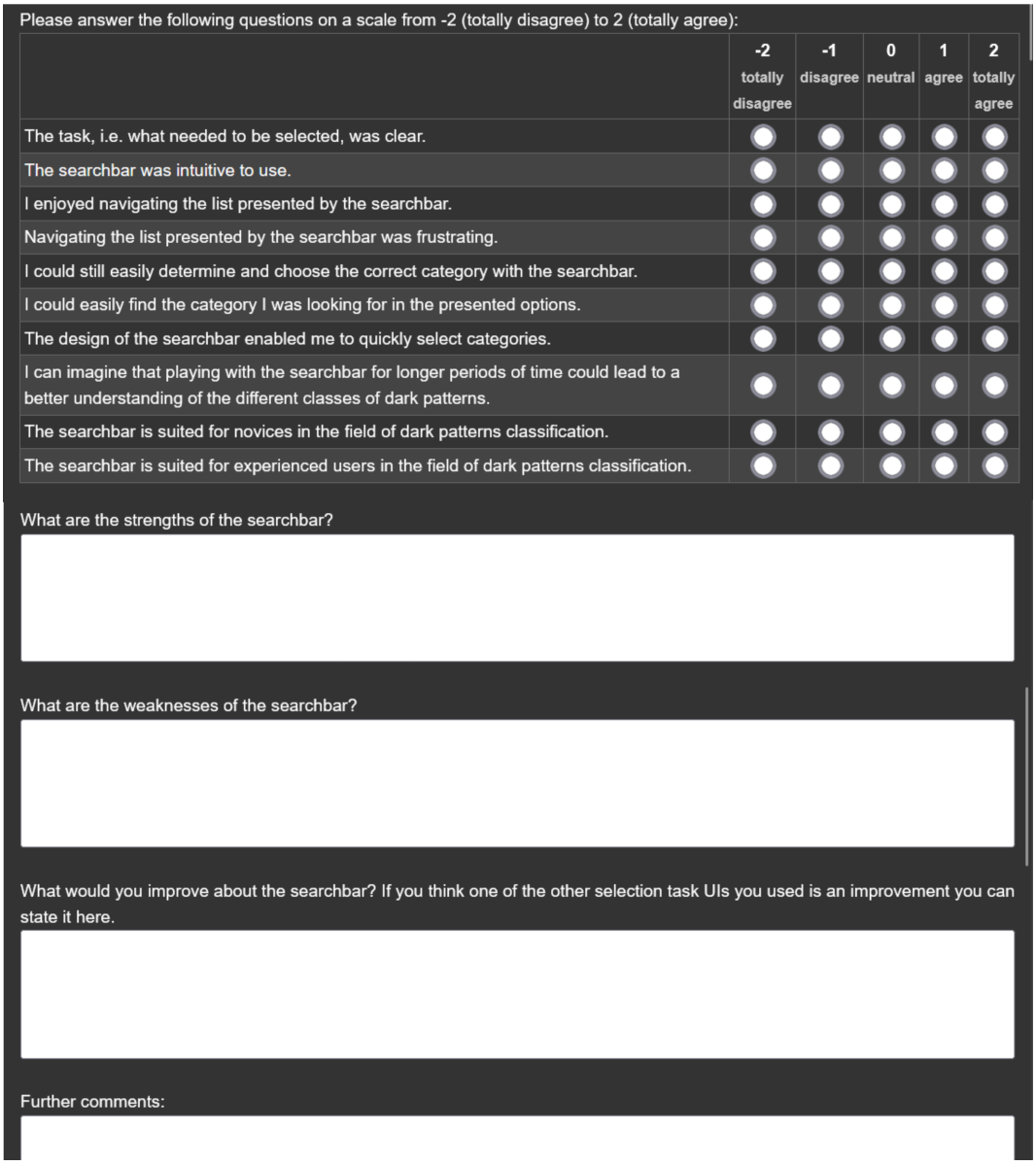
(b) The first part of the second questionnaire asked for the participants' opinions about the different UI styles (*Sidebar* vs. *Bubbles*) they interacted with. The questions in the figure were asked 2 times in total, once for each UI style (replacing "sidebar" with "bubbles" and vice versa), along with a corresponding picture (from Chapter 3.2.1, "Selection Tasks & Interface Styles") of an interface in the corresponding style of either the *Sidebar* or the *Bubbles* as a reminder.

Figure B.2: The three questionnaires included in our user study (cont.)



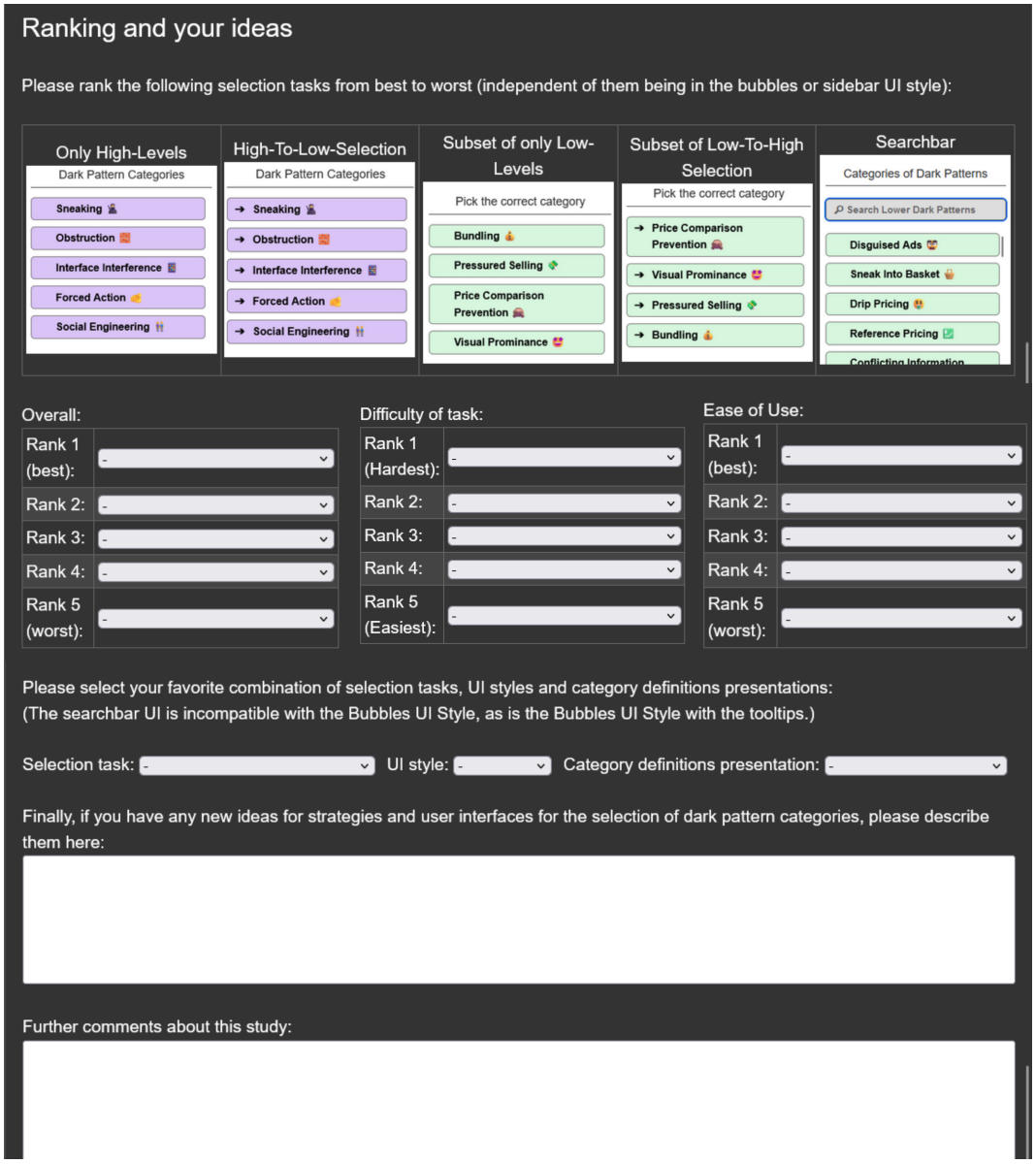
(c) The second part of the second questionnaire asked for the participants' opinions about the category definition presentations (*Tooltips* vs. *Book of Dark Patterns*) they interacted with. The participants could choose to skip the questions for the definitions representations if they didn't use one of them, or both. The questions in the figure were asked 2 times in total, once for each definitions representation (replacing "tooltips" with "book of dark patterns" and vice versa), along with a corresponding picture (from Chapter 3.2.1, "Selection Tasks & Interface Styles") of either the *Tooltips* or the *Book* as a reminder.

Figure B.2: The three questionnaires included in our user study (cont.)



(d) The first part of the third questionnaire asked for the participants' opinions about the searchbar UI they interacted with, along with a corresponding picture (from Chapter 3.2.1, "Selection Tasks & Interface Styles").

Figure B.2: The three questionnaires included in our user study (cont.)



(e) The second part of the third questionnaire asked for the participants' final rankings of the selection tasks (including the *Searchbar* UI), independent from the interface style, along with corresponding pictures (from Chapter 3.2.1, "Selection Tasks & Interface Styles"). If the participants had the *Bubbles* style for the first 4 interfaces, images of those were presented instead of the *Sidebar* ones in the figure. Further, the participants were asked to select their favorite (possible) combination of selection tasks, interface styles and definition presentations. Finally, they were asked for any own ideas for strategies and menu interfaces for the classification of dark pattern categories.

Figure B.2: The three questionnaires included in our user study (cont.)

Bibliography

[1] Sanju Ahuja and Jyoti Kumar. Understanding Ethical Thinking in Design Education: A Linkographic Study. In *EduCHI '24: Proceedings of the 6th Annual Symposium on HCI Education*, EduCHI '24, New York, NY, USA, 2024. Association for Computing Machinery. doi.org/10.1145/3658619.3658620.

[2] Benjamin B. Bederson. Fisheye menus. In *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology*, UIST '00, page 217–225, New York, NY, USA, 2000. Association for Computing Machinery. doi.org/10.1145/354401.354782.

[3] Aditi Bhoot, Mayuri Shinde, and Wricha Mishra. Towards the Identification of Dark Patterns: An Analysis Based on End-User Reactions. In *IndiaHCI 2020: 11th Indian Conference on Human-Computer Interaction*, pages 24–33, 11 2020. doi.org/10.1145/3429290.3429293.

[4] Kerstin Bongard-Blanchy, Arianna Rossi, Salvador Rivas, Sophie Doublet, Vincent Koenig, and Gabriele Lenzini. "I am Definitely Manipulated, Even When I am Aware of it. It's Ridiculous!" - Dark Patterns from the End-User Perspective. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference*, DIS '21, page 763–776, New York, NY, USA, 2021. Association for Computing Machinery. doi.org/10.1145/3461778.3462086.

[5] Jan O. Borchers. A pattern approach to interaction design. In *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, DIS '00, page 369–378, New York, NY, USA, 2000. Association for Computing Machinery. doi.org/10.1145/347642.347795.

[6] James V. Bradley. Complete Counterbalancing of Immediate Sequential Effects in a Latin Square Design. *Journal of the American Statistical Association*, 53(282): 525–528, 1958. doi.org/10.1080/01621459.1958.10501456.

[7] Harry Brignull, Mark Leiser, Christina Santos, and Kosha Doshi. Deceptive patterns – user interfaces designed to trick you, 2023. URL https://www.deceptive.design/.

[8] Raluca Budiu. Search Is Not Enough: Synergy Between Navigation and Search, 2014. URL https://www.nngroup.com/articles/search-not-enough/.

[9] Christoph Bösch, Benjamin Erb, Frank Kargl, Henning Kopp, and Stefan Pfattheicher. Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns. *Proceedings on Privacy Enhancing Technologies*, 2016:237–254, 07 2016. doi.org/10.1515/popets-2016-0038.

[10] J. Callahan, D. Hopkins, M. Weiser, and B. Shneiderman. An empirical comparison of pie vs. linear menus. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '88, page 95–100, New York, NY, USA, 1988. Association for Computing Machinery. doi.org/10.1145/57167.57182.

[11] Evan Caragay, Katherine Xiong, Jonathan Zong, and Daniel Jackson. Beyond Dark Patterns: A Concept-Based Framework for Ethical Software Design. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, CHI '24, page 1–16. ACM, May 2024. doi.org/10.1145/3613904.3642781.

[12] Andy Cockburn and Andrew Gin. Faster cascading menu selections with enlarged activation areas. In *Proceedings of Graphics Interface 2006*, GI 2006, pages 65–71, Toronto, Ontario, Canada, 2006. Canadian Human-Computer Communications Society.

[13] Gregory Conti and Edward Sobiesk. Malicious interface design: exploiting the user. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, page 271–280, New York, NY, USA, 2010. Association for Computing Machinery. doi.org/10.1145/1772690.1772719.

[14] Mary Dondlinger. Educational Video Game Design: A Review of the Literature. *Journal of Applied Educational Technology*, 4, 01 2007.

[15] Ralf Dörner, Stefan Göbel, Wolfgang Effelsberg, and Josef Wiemeyer. *Serious Games - Foundations, Concepts and Practice*. Springer Cham, 2016. doi.org/10.1007/978-3-319-40612-1.

[16] Kevin Fiedler. Exploring Game Mechanics for the Design of Dark Pattern Learning Games. Master's thesis, RWTH Aachen University, 2024.

[17] BJ Fogg. A behavior model for persuasive design. In *Persuasive '09*, New York, NY, USA, 2009. Association for Computing Machinery. doi.org/10.1145/1541948.1541999.

[18] James Paul Gee. What video games have to teach us about learning and literacy. *Comput. Entertain.*, 1(1):20, October 2003. doi.org/10.1145/950566.950595.

[19] Colin Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin Toombs. The Dark (Patterns) Side of UX Design. In *CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1 – 14, 04 2018. doi.org/10.1145/3173574.3174108.

[20] Colin M. Gray, Cristiana Santos, and Nataliia Bielova. Towards a Preliminary Ontology of Dark Patterns Knowledge. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI EA '23, New York, NY, USA, 2023. Association for Computing Machinery. doi.org/10.1145/3544549.3585676.

[21] Colin M. Gray, Johanna T. Gunawan, René Schäfer, Nataliia Bielova, Lorena Sanchez Chamorro, Katie Seaborn, Thomas Mildner, and Hauke Sandhaus. Mobilizing Research and Regulatory Action on Dark Patterns and Deceptive Design Practices. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '24, New York, NY, USA, 2024. Association for Computing Machinery. doi.org/10.1145/3613905.3636310.

[22] Colin M. Gray, Cristiana Teixeira Santos, Nataliia Bielova, and Thomas Mildner. An Ontology of Dark Patterns Knowledge: Foundations, Definitions, and a Pathway for Shared Knowledge-Building. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA, 2024. Association for Computing Machinery. doi.org/10.1145/3613904.3642436.

[23] Mark Griffiths. The educational benefits of videogames. *Education and Health*, 20:47–51, 01 2002.

[24] Johanna Gunawan, Cristiana Santos, and Irene Kamara. Redress for Dark Patterns Privacy Harms? A Case Study on Consent Interactions. In *Proceedings of the 2022 Symposium on Computer Science and Law*, CSLAW '22, page 181–194, New York, NY, USA, 2022. Association for Computing Machinery. doi.org/10.1145/3511265.3550448.

[25] Harry Hochheiser and Ben Shneiderman. Performance Benefits of Simultaneous Over Sequential Menus as Task Complexity Increases. *Int. J. Hum. Comput. Interaction*, 12:173–192, 06 2000. doi.org/10.1207/S15327590IJHC1202_2.

[26] David R. Krathwohl. A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, 41(4):212–218, 2002. doi.org/10.1207/s15430421tip4104_2.

[27] K. Kronhardt and J. Gerken. Start playing around: serious & persuasive games as a viable counter-measure against deceptive patterns?". In *Proceedings of the Workshop Mobilizing Research and Regulatory Action on Dark Patterns and Deceptive Design Practices (DDPCHI 2024)*, 2024.

[28] Yuwen Lu, Chao Zhang, Yuewen Yang, Yaxing Yao, and Toby Jia-Jun Li. From Awareness to Action: Exploring End-User Empowerment Interventions for Dark Patterns in UX. *Proceedings of the ACM on Human-Computer Interaction*, 8 (CSCW1), apr 2024. doi.org/10.1145/3637336.

[29] Jamie Luguri and Lior Jacob Strahilevitz. Shining a Light on Dark Patterns. *Journal of Legal Analysis*, 13(1):43–109, 03 2021. doi.org/10.1093/jla/laaa006.

[30] Deborah Löschner and Sebastian Pannasch. Measuring the Deceptive Potential of Design Patterns: A Decision-Making Game. In *Proceedings of the Workshop Mobilizing Research and Regulatory Action on Dark Patterns and Deceptive Design Practices (DDPCHI 2024)*, 07 2024.

[31] Maximilian Maier and Rikard Harr. Dark design patterns: An end-user perspective. *Human Technology*, 16(2):170–199, Aug. 2020.

[32] Arunesh Mathur, Gunes Acar, Michael Friedman, Elena Lucherini, Jonathan R. Mayer, Marshini Chetty, and Arvind Narayanan. Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. *CoRR*, abs/1907.07032, 2019.

[33] Arunesh Mathur, Mihir Kshirsagar, and Jonathan Mayer. What Makes a Dark Pattern... Dark? Design Attributes, Normative Considerations, and Measurement Methods. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery. doi.org/10.1145/3411764.3445610.

[34] Thomas Mildner, Gian-Luca Savino, Philip R. Doyle, Benjamin R. Cowan, and Rainer Malaka. About Engaging and Governing Strategies: A Thematic Analysis of Dark Patterns in Social Networking Services. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, page 1–15. ACM, April 2023. doi.org/10.1145/3544548.3580695.

[35] Emily Naul and Min Liu. Why Story Matters: A Review of Narrative in Serious Games. *Journal of Educational Computing Research*, 58:073563311985990, 07 2019. doi.org/10.1177/0735633119859904.

[36] Chinenye Ndulue and Rita Orji. Games for Change - A Comparative Systematic Review of Persuasive Strategies in Games for Behaviour Change. *IEEE Transactions on Games*, pages 1–1, 01 2022. doi.org/10.1109/TG.2022.3159090.

[37] Kent L. Norman. Better Design of Menu Selection Systems Through Cognitive Psychology and Human Factors. *Human Factors*, 50(3):556–559, 2008. doi.org/10.1518/001872008X288411.

[38] Harri Oinas-Kukkonen and Marja Harjumaa. Persuasive Systems Design: Key Issues, Process Model, and System Features. *Communications of the Association for Information Systems*, 24, 03 2009. doi.org/10.17705/1CAIS.02428.

[39] Jan Plass, Bruce Homer, and Charles Kinzer. Foundations of Game-Based Learning. *Educational Psychologist*, 50:258–283, 10 2015. doi.org/10.1080/00461520.2015.1122533.

[40] René Röpke, Vincent Drury, Ulrike Meyer, and Ulrik Schroeder. Exploring and Evaluating Different Game Mechanics for Anti-Phishing Learning Games. *International Journal of Serious Games*, 9:23–41, 09 2022. doi.org/10.17083/ijsg.v9i3.501.

[41] Nabil Saleh, Jon Roozenbeek, Fadi Makki, William McClanahan, and Sander van der Linden. Active inoculation boosts attitudinal resistance against extremist persuasion techniques: a novel approach towards the prevention of violent extremism. *Behavioural Public Policy*, 8:1–24, 02 2021. doi.org/10.1017/bpp.2020.60.

[42] Krystian Samp and Stefan Decker. Supporting menu design with radial layouts. In *Proceedings of the International Conference on Advanced Visual Interfaces*, AVI '10, page 155–162, New York, NY, USA, 2010. Association for Computing Machinery. doi.org/10.1145/1842993.1843021.

[43] René Schäfer, Paul Miles Preuschoff, and Jan Borchers. Investigating Visual Countermeasures Against Dark Patterns in User Interfaces. In *Proceedings of Mensch Und Computer 2023*, MuC '23, page 161–172, New York, NY, USA, 2023. Association for Computing Machinery. doi.org/10.1145/3603555.3603563.

[44] René Schäfer, Paul Miles Preuschoff, René Röpke, Sarah Sahabi, and Jan Borchers. Fighting Malicious Designs: Towards Visual Countermeasures Against Dark Patterns. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA, 2024. Association for Computing Machinery. doi.org/10.1145/3613904.3642661.

[45] Steve Sheng, Bryant Magnien, Ponnurangam Kumaraguru, Alessandro Acquisti, Lorrie Cranor, Jason Hong, and Elizabeth Nunge. Anti-Phishing Phil: The Design and Evaluation of a Game That Teaches People Not to Fall for Phish. In *ACM International Conference Proceeding Series*, volume 229, pages 88–99, 07 2007. doi.org/10.1145/1280680.1280692.

[46] Similarweb, 2024. URL https://www.similarweb.com/top-websites/.

[47] Sponge B. Squarepants, Squidward J. Q. Tentacles, and Eugene Harold Krabs. How to Make the Perfect Krabby Patty: A Step by Step Guide for Novel Fry Cooks. In *Proceedings of the 1999 Underwater Conference on Human Factors in Food Preparation*, Bikini Bottom, 5 1999. Association for Aquatic Marinade.

[48] Stephen Tang, Martin Hanneghan, and A. El Rhalibi. Introduction to Games-Based Learning. *Games-Based Learning Advancements for Multi-Sensory Human*

Computer Interfaces: Techniques and Effective Practices, pages 1–17, 01 2009. doi.org/10.4018/978-1-60566-360-9.ch001.

[49] Ingvar Tjostheim, Vanessa Ayres-Pereira, Chris Wales, Angela Manna, and Simon Egenfeldt-Nielsen. Dark Pattern: A Serious Game for Learning About the Dangers of Sharing Data. *Proceedings of the 16th European Conference on Games Based Learning*, 16:774–783, 10 2022. doi.org/10.34190/ecgbl.16.1.872.

[50] Zhonggen Yu. A Meta-Analysis of Use of Serious Games in Education over a Decade. *International Journal of Computer Games Technology*, 2019:1–8, 02 2019. doi.org/10.1155/2019/4797032.

[51] José Pablo Zagal, Staffan Björk, and Chris Lewis. Dark patterns in the design of games. In *International Conference on Foundations of Digital Games*, 2013.

[52] Panayiotis Zaphiris, Ben Shneiderman, and Kent Norman. Expandable Indexes Versus Sequential Menus for Searching Hierarchies on the World Wide Web. *Behaviour and Information Technology*, 08 2001. doi.org/10.1080/01449290210000090045.

Index

Typeset November 26, 2024