Understanding the Impact of Dark Pattern Detection on Online Users

Ryan Wood

Thesis submitted to the Faculty of
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Science & Application

Dwayne C. Brown, Chair
Daniel R. Dunlap
Donald S. McCrickard

May 8, 2023
Blacksburg, Virginia

Keywords: Dark Patterns, Deceptive Design, User Study

Understanding the Impact of Dark Pattern Detection on Online Users

Ryan Wood

(ABSTRACT)

Dark Patterns are a variety of different software designs that are used to manipulate and mislead the users of an application or service. These patterns range from making it harder to end a subscription service, adding additional charges to a purchase, or having the user give out data or personal information. With how widespread and varied dark patterns are, it led to us creating a way to detect and warn users of different dark patterns.

In this study, we created Dark Pattern Detector, a Chrome extension that would help users detect and understand three different dark patterns: Hidden Costs, Disguised Ads, and Sneak into Basket. This extension was made to detect each of these patterns on any web page while not requiring any information from the user or their data. Study participants installed the extension and completed a series of tasks given to them that would occur on different websites containing the previous dark patterns. After completing the tasks, the users were surveyed to give feedback on what they thought of the extension and what suggestions for change they had.

In the study, we had 40 participants and we found that 50% of the users were completely unfamiliar with dark patterns and that 77.5% have used extensions before. For the five tasks, each one had a majority of the participants successfully complete them. Finally, when asked about what they thought, the majority of the participants gave positive feedback claiming that they found the extension useful, interesting, and a good idea. Many participants also gave useful feedback about what changes or additions they would like to see. With our

results, we can help users have a better understanding of dark patterns and have created a baseline for any future research done on dark pattern knowledge and detection.

Understanding the Impact of Dark Pattern Detection on Online Users

Ryan Wood

(GENERAL AUDIENCE ABSTRACT)

Dark patterns are designs on the internet that websites use to trick its users. They may be used to hide advertisements, make the user spend more time or money on their website or more. Our goal was to create a way to help protect anyone on the internet and their information.

For this study, we created a program called Dark Pattern Detector that would help the users see different dark patterns that appeared on websites. A study was conducted that had the participants use our program and give us feedback on what they thought of it as well as data on how well it worked. Out of the 40 participants, we found that half the users were unfamiliar with what dark patterns were. Once they completed the study, we saw that the majority of users were able to complete tasks while using our program and gave positive feedback.

Seeing the positive feedback and results from our study, we believe that we can help users not get tricked by these patterns and help forward future research on Dark Patterns.

Acknowledgments

I would like to thank my advisor, Dr. Brown for working with me throughout my research and being extremely helpful for the entire process. I would not have been able finish everything without his help.

I would also like to thank my committee members, Dr. Dunlap and Dr. McCrickard whom I appreciated having them for help and feedback and were a pleasure to work with.

Next, I want to thank my research group for helping and giving feedback to help prepare for my defense and for general feedback on my project overall

Finally, I want to thank my friends, family, and girlfriend for being there to support me through my thesis, for listening to me talk about this for months, and for making sure I ate and slept.

Contents

List of Figures

List of Tables

List of Abbreviations

DOM Document Object Model

RQ Research Question

RQ is for Research Questions and represents the main questions that the study was interested in. DOM is for Document Object Model which is the structure of the document (webpages) and connects them to the scripts that represent them

Chapter 1

Introduction

Dark patterns are tricks and traps that have been made and added into websites and apps to try and mislead or coerce the user into doing something that benefits the maker of the pattern and potentially hurts the user. They take advantage of us by making advertisements appear to be part of the article or having a fake download or by hiding important information from the user itself. Some examples of dark patterns are: Disguised Ads -which are advertisements that are hidden on a website to make the user believe they are a part of it, Confirmshaming - which is when the user is guilted into doing something they may not have previously done, Sneak into Basket - where websites add additional items to the cart to have the users spend more if they don't notice, Preselection - which selects an option for the user in attempt to influence their decision making, Hidden Costs/Fees - which keeps prices hidden until the checkout screen to have the user not notice the increase or not care to change it, and Roach Motel/Hard to Cancel - which is when services make it hard to cancel or unsubscribe [3]. Studies have been done to find out how common they are and how they have been used to trick people. One study done by Di Geronimo et al. found that 95% (out of 240 popular mobile applications) had at least one dark pattern and Mathur et al. found that 11% of 11k e-commerce websites also used Dark Patterns [8][12].

We decided as a way to help users understand dark patterns more and help make them more visible on the web, we made the Dark Pattern Detector. The Dark Pattern Detector

is a browser extension that we published to the Chrome web store that could automatically detect a set of dark patterns. The three patterns that we choose to detect were Disguised Ads, Sneak into Basket, and Hidden Costs/Fees.

For our study, we explored the four following Research Questions (RQ)s:

RQ1 How familiar are users with the concept of Dark Patterns? Dark Patterns are found in many places around the web but outside of those who work on them or with them, they are not as commonly known. We want to see how many people actually know what Dark Patterns are and how familiar they are with them. Through this question, we can get a better understanding of how much we need to do in order to spread awareness to combat Dark Patterns.

RQ2 How can we make dark patterns more visible on the internet? One of the main issues with dark patterns is that even if someone knows about them, they can still be fooled if they have no way to see them. We want to answer this to see how familiar people are with extensions and if they could be used to spread information about dark patterns.

RQ3 What ways can users potentially avoid or block dark patterns? After knowing about dark patterns, the next step is knowing how to avoid them. We have created tasks to see if using the extension would help the study participant's ability to find and avoid dark patterns.

RQ4 How usable do users find the Dark Pattern Detector? Our final question is about how useful the participants found the extension and if there were any issues with it. This covered if the extension did or did not work properly and if the users thought it was

useful and aesthetically appealing.

With these questions in mind, we created a user survey along with the extension as a way to generate feedback from users and to test how well they were able to navigate specific websites that contained dark patterns when they were using the extension. The survey gave them tasks to complete that were measured for correctness and different feedback questions about dark patterns, browser extensions, and what the users thought about the Dark Pattern Detector.

Once completed, we had 40 unique participants finish the survey. Out of those participants, it was found that 20 or 50.0% were not familiar with dark patterns at all, and out of the 20 who were, six were only somewhat familiar with the concept or the phrase "Dark Patterns" (RQ1). The majority of the participants were able to successfully complete all of the given tasks and found the extension to be useful overall. The tasks were split into disguised ads tasks which would have the users navigate a website that has hidden advertisements with the extension on and notification tasks which were for Sneak into Basket and Hidden Fees and would have the user complete an action and look for a notification from the extension. For the disguised ads tasks, Tasks 1-3 had users successfully complete the task at 87.5%, 90%, and 92.5% respectively. For the notification tasks, 72.5% claimed to have the notification appear for Task 4 and 82.5% for Task 5. The feedback that was given was mostly positive with the main complaint being that the yellow highlight for the advertisements was too bright. However, this was not unanimous as other participants found it useful to be as bright as it was. Overall, we found there was mostly positive feedback and success from participants in completing the tasks showing that our study was useful in helping dark pattern detection as well as creating a strong baseline for future studies to build off of.

Chapter 2

Background

In this section, we show examples of different dark patterns that are used around the web and show how they can be used to manipulate someone. To start off, we show an example of the Sneak into Basket on the website GoDaddy.com [10]. In Figure 2.1, we see different prices for different domain registrations. For this example, we look at the .net registration which claims to be $14.99 from $22.99. The next figure, Figure 2.2, shows the next step of

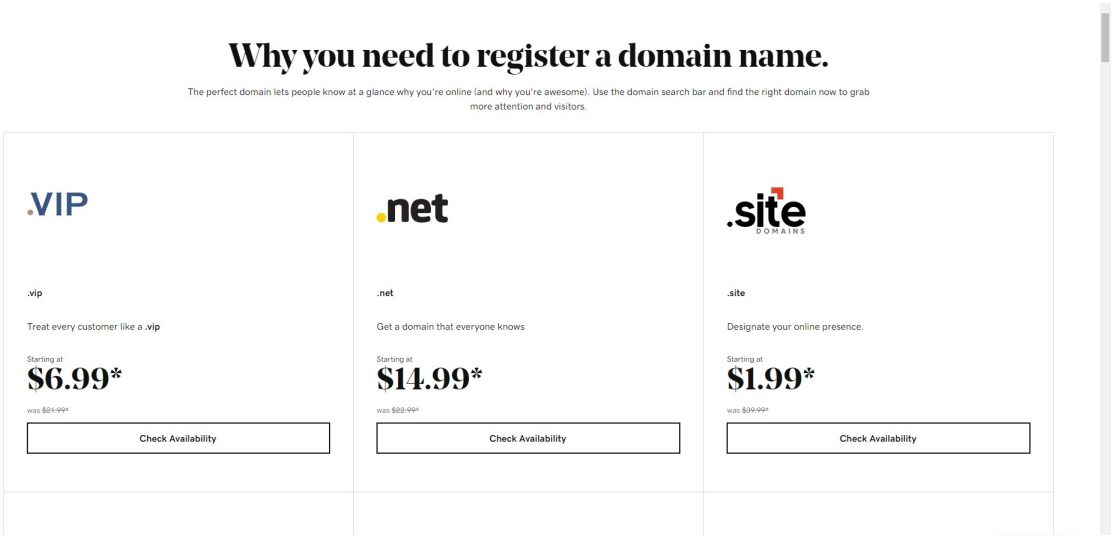

Figure 2.1: The home page of GoDaddy.com

the processing with choosing the domain name and includes additional options and a statement that the reduced price is only applicable for the first year of a two-year registration. This also borders on being a Hidden Cost dark pattern as well by having additional fees that can be missed by the user. Figure 2.3 shows an automatically selected option that is also

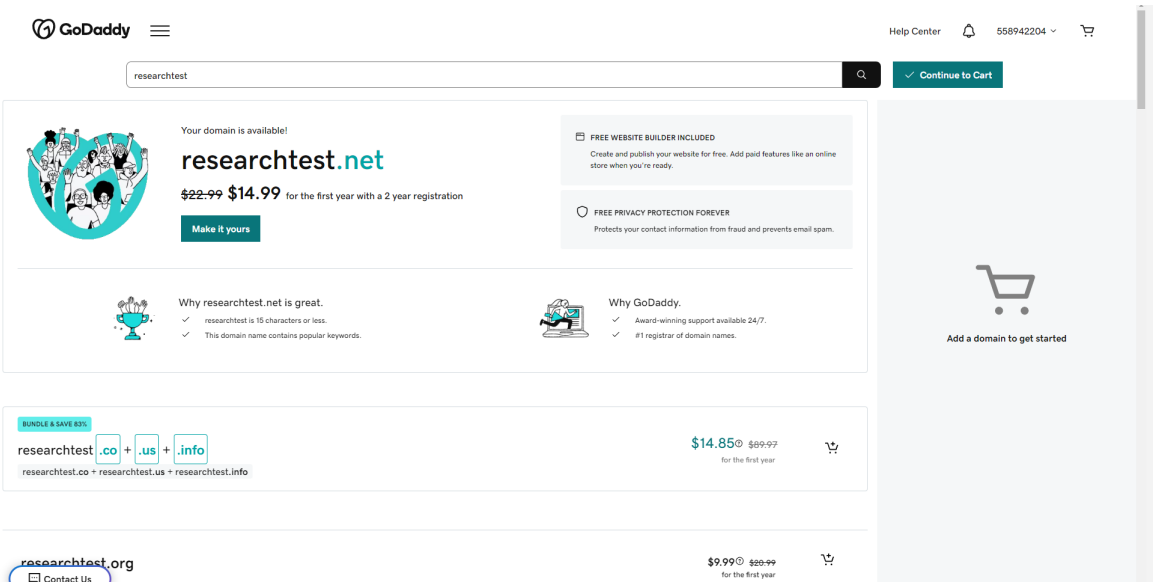

Figure 2.2: Registration page for choosing a name and domain type.

recommended by the site for domain protection and finally, Figure 2.4 shows the final cart which has a final total of $57.96. The Sneak into Basket pattern is shown by the automatic selection of the domain protection. Without clicking on anything outside of the Continue to Cart button, the price has jumped significantly from what was expected from the initial figure. Additionally, other dark patterns are also partially shown here as well. The first is Hidden Costs coming from the prices being shown for one year but the registration is for two years causing an additional spike in the price for the registration and domain protection. Misdirection is also partially seen by the site trying to distract the user with all of the crossed-through prices and notes on how much they are saving while the user is paying more than the expected price.

Another example of dark patterns around the web is seen in Figure 2.5. This image is taken from Gmail, a popular e-mail service run by Google. In the figure, you see the website telling the user about the newest version of the Gmail app for iPhone and how it is smarter and faster than what they are currently using. They then give two options, the confirm

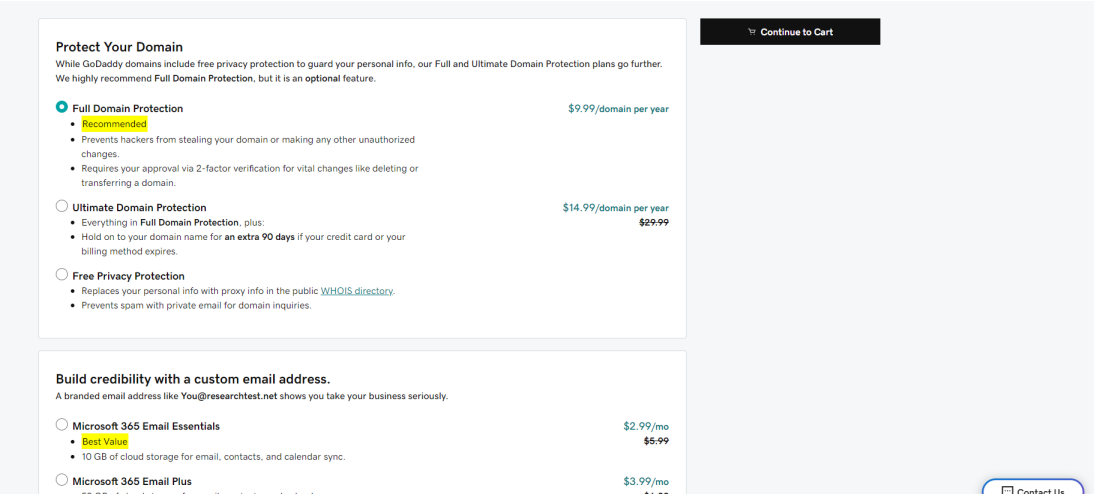

Figure 2.3: Domain Protection Page where you can choose ad ons

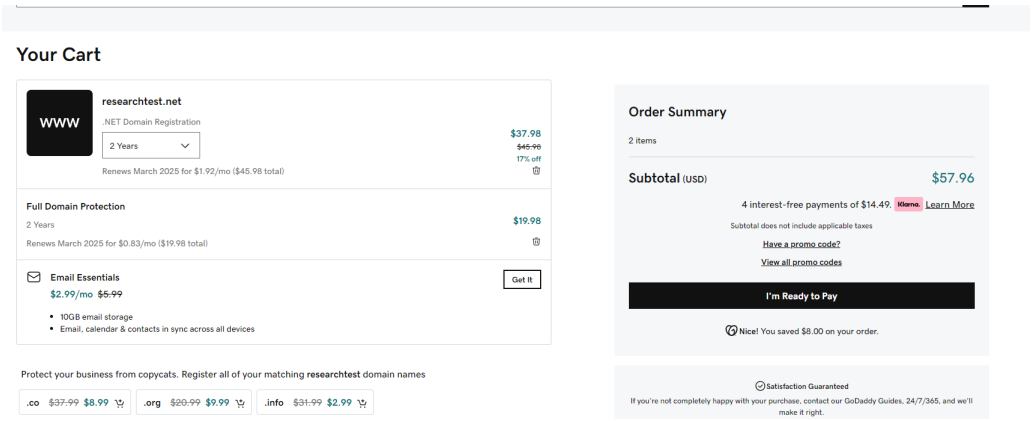

Figure 2.4: Checkout page for godaddy.com

option being a blue button with white text with the denial option being blue text with no button that says "I don't want smarter email". This is an example of Confirmshaming which is when a service tries to guilt a user into picking a specific option and agreeing to something by making the denial option seem like a bad thing and shame them.

Both of these examples are from popular websites with GoDaddy having 20.9 million customers in 2022 [sec.gov site] and Gmail having around 1.8 billion active users globally in 2023 [16]. While both of these examples are not life-altering if fallen for, they still are there

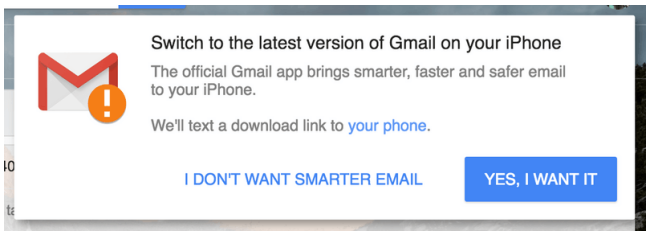

Figure 2.5: Popup on an iPhone browser [3]

to trick the user into benefitting the company. What we want to do is find ways to help prevent users from being tricked and to be more informed on dark patterns and be able to detect and avoid them.

Chapter 3

Review of Literature

In our literature review section, we looked at different studies and resources on dark patterns and how they are defined, detecting dark patterns, and studies surveying users involving dark patterns.

For looking deeper into dark patterns, we started with Harry Brignull's et al. site "deceptive.design". Brignull et al. created a collection of the different types of dark patterns and examples of each kind in the real world [3]. They defined dark patterns as "tricks used in websites and apps that make you do things you did not mean to such as buying or signing up for something" and had a list of 16 different types of dark patterns such as the ones mentioned prior [3]. We were able to go through this website to get a better understanding of what dark patterns are and how they are displayed around the internet. Additionally, they have also keep a list of different laws and judicial cases from around the world that are relevant to dark patterns and deceptive designs [3]. Another study that references Brignull's work was by Mathur et al. that discusses what makes a dark pattern "dark" [13]. In this study, they created a list of literature that references and defines dark patterns and found that while there was variance in how everyone defined them [13]. For example, while Brignull said the required characteristic of dark patterns in User Interface was "Trickery", Conti & Sobiesk said they had to be "Malicious" and Bosch et al. described it as misleading [2, 3, 6, 13].

Studies also described them as ”steering”, ”seductive”, and ”coercive” and multiple studies including Brignull and Bosch et al. say that dark patterns effect is to ”subvert user intent” [2, 3, 13]. Another study done by Cara et al. created another list of dark patterns and what they were and added a few additional types such as Infinite Scroll which is a way to manipulate the user to stay on the service since there is seemingly endless content, Pseudo Currency in an attempt to have the user forget that the money on their service is real money with another name, and also contained patterns that were the same as the one's on Brignull's list with a different name (Confirmshaming = Guilt Tripping) [**?**]. With how many types there are and the variations in their descriptions and names, we want to try and help give users a better understanding of what dark patterns are and how they appear on the web.

For detecting dark patterns, we started with a series of papers that looked into cookie consent decisions and detecting dark patterns in them. The first paper was ”Circumvention by Design - dark patterns in cookie consent for online news outlets” by Soe et. al. in which they manually went through the consent notices for cookies of 300 different news websites [17]. They found that of the 300, 297 were said to use some form of dark pattern in trying to obtain the user's consent. However, when another study was done by Soe et. al as a way to use machine learning and automatically detect dark patterns, they were unable to successfully create a tool that was able to autonomously detect the dark patterns without additional studies and manual help [18]. While not for as large of a data set, our goal was to be able to detect dark patterns completely autonomously. Another study by Hausner et. al. was completed on detecting a cookie banner and changing the Document Object Model (DOM) and was able to surround and add borders to buttons on the banner through mostly automatic means [11]. In our study, we are also looking at modifying the DOM as a way to change the webpage and detect and highlight the dark patterns.

Another study that we looked at was "Smart Dark Pattern Detection: Making Aware of Misleading Patterns Through the Intended App" by Raju et. al. In this study, the authors created an app that could automatically detect dark patterns by putting websites into the app and parsing through the page, and comparing it to pre-made signatures for a list of dark patterns and created a report that explained what it found [15]. While they were able to successfully detect dark patterns autonomously in a novel way, we tried to create a method that would not require specific signatures that would have to be added and it happen automatically on the page instead of through an app and a report. We also looked at a study done by Curley et. al. that created a framework and broke down dark patterns into a list of whether or not they could be detected automatically, manually, or at all [7]. They determined that five different types of dark patterns could not be detected at all due to their variance and only the Roach Motel dark pattern, a pattern that describes services that are intentionally difficult to get out of, could be detected fully automatically [7][3].

Additionally, Geronimo et. al. also surveyed 589 individuals by having them go through a few applications and being asked to identify dark patterns in them and it was found that 55% of the participants did not spot the dark patterns, 20% were unsure, and the remaining 25% were able to spot at least one [8]. Similarly, the study "Dark Patterns from the End-User Perspective" by Bongard-Blanchy et. al. surveyed 406 participants to study how people responded to dark patterns [1]. They found that there was an inverse correlation between dark pattern recognition and the likelihood to be influenced by a dark pattern [1]. With these previous studies in mind, we wanted to create a way to autonomously detect dark patterns that would help show the user what they were looking at and teach them more

about dark patterns.

Chapter 4

Dark Pattern Detector

The process for creating the Dark Pattern Detector was broken down into the original design of the tool, or what we wanted the tool to do, and the practical implementation, or what we were able to implement for this study.

The original plan of the Dark Pattern Detector was to create a tool that could detect a set of dark patterns, could be distributed easily, and could inform the user more about dark patterns both through giving access to resources and through making them more visible (the detection). The set of dark patterns was determined by the us going through Harry Brignull's website, "deceptive.design" which contains a comprehensive list of Dark Pattern types, examples of each of the types, and additional information about what dark patterns are [3]. To choose the dark patterns, we wanted to pick types that were reasonable to automatically detect and did not make user's meet a requirement by having something before the study (such as having a subscription that was difficult to unsubscribe from for Roach Motel). The first pattern picked was Disguised Advertisements as it was a commonly occurring pattern as almost every website has some form of advertisements (hidden or not) and did not require the user to do anything outside of go to a website to test the extension. The next patterns picked were Sneak into Basket and Hidden Fees which both happened near the end of purchasing an item or service but did not require any payment from the user. For each of these patterns, we wanted to have the users see them on real websites across the web and have them be highlighted so that the user could easily see what was happening and

how the website was trying to trick or persuade them. To be able to distribute the tool, we wanted to create a program that did not require much user interaction so that anyone could easily access the program without having to go through an arduous, technical process to use it. This led to the idea of using a browser extension that could be published and downloaded by the participants of the study. To give the users information and resources about dark patterns, we wanted to have the Dark Pattern Detector explain what was happening when a pattern was detected and then give access to examples and links to websites that added additional, in-depth information about the pattern.

To detect Disguised Ads, we decided that we wanted to modify the webpage through the DOM to make the hidden advertisements more visible. To do so, we check element tags such as “id”, “class”, and “iframe” if it was an image and search for instances of ad in any of the tags. If found, the source of the potential advertisement was checked against the location of the current webpage to see if they were from the same site. For example, if the user was on YouTube, another video that coincidentally met the criteria for containing ad but were still on YouTube would not be flagged but an ad from GoogleAdSense would be flagged. Once flagged, the extension would then either modify the background color of the element containing the ad or, if the advertisement was an image, would hide it entirely. The images were hidden as they proved to be resistant to changing the color and other types of modification as it would change the page behind the image and would not affect the image itself. Since some disguised ads come in the form of a fake download button image, we still found it important to modify them for the purpose of the extension. For the color change, we decided on using a bright yellow for the background color with a thick black border around the edges. The choice for yellow came from needing to find a color that black text showed up well on and was not too light or pale that it was difficult to see. The black border was added as a way to prevent any users that had difficulty seeing the color from missing it or for it blending in

on any websites or advertisements that used a similar background color. For the notification system, we wanted to create a way to have the extension alert the user before they would have to pay anything but after something had been snuck into the basket or a hidden fee was added to their total. To do this, we used "alert" which would create a pop-up at the top, middle of the window that we could add a custom message to. We originally wanted the notification to appear when the price changed for Hidden Costs and when the total items in the basket changed either by more than the user intended to add at once or if it changed when the user had not explicitly added anything. However, due to difficulty that came with tracking these changes, we decided to change it so that the notification would appear on the checkout page when the price changes for both and gave a warning to the user stating "The cost of your purchase may have gone up due to hidden fees and items being snuck into your basket. Check out 'deceptive.design' to learn more about these patterns." This alert would appear with a singular button that says "OK" and would stay until the user interacted with it. Our notification would appear once the page that was flagged finished loading.

For distributing the browser extension, we decided that we were going to publish the extension on the Chrome web store. While Google Chrome is a commonly used browser, it did create a restriction that only people with the browser could complete the study. This came from different browsers having different requirements and formats for publishing an extension and we wanted to keep the extension consistent for all the participants. To publish the web store, we had to first create and pay a $5 fee to create a developer account. This would allow us to attempt to publish the Dark Pattern Detector. To publish an extension, it has to be approved by the web store an issue that came with that was that if the extension required any permission or stored user data, it would significantly increase the amount of approval/rejection time for the extension. To deal with this we used as few permissions as possible and only looked at data from the current tab the user was on and did not carry it

over or store it. While this led to some of the issues with the notification system, we were able to continue using alert for our notification.

Chapter 5

Methodology

For this study, we first created a Google Chrome extension to help us detect the chosen dark patterns. Next, we created a user survey that contained tasks for the participants to complete to help us understand their familiarity with dark patterns, how well they can be detected, and to collect feedback.

5.1 User Survery

The user survey was made with the research questions in mind for both the questions and the tasks. First, demographic information was collected from the participant to be able to confirm that each response was from a unique participant. While this data was collected, all participants are to remain anonymous and will be referred to by the number of their submissions, ex. "Participant 13". After the demographic questions, the participants were asked if they knew about dark patterns and if so, how much they were familiar with them. This question was to help answer RQ1 by seeing how many people are aware of what dark patterns are. Next, we gave a definition for dark patterns along with an explanation of what the participants would be doing. In the explanation we state what types of dark patterns we will be looking at and that they will be downloading a Chrome extension and completing tasks. If the user was not willing or able to participate in the study after reading this, they were able to close the study as an opt-out. We then ask the participant ask for how often

and how familiar they are with browser extensions. Additionally, we ask if there were any issues installing the extension when they are prompted to. Both of these questions are for collecting information for RQ2 on participants' experience and understanding of extensions.

5.1.1 Pre-Tasks

Initially, the participants were asked to complete three preemptive tasks (shortened to pre-tasks). The first two involved them looking at the websites "softpedia" and "dafont" both of which are used in later tasks and have many advertisements on them. Softpedia is a popular software downloading site used for Macs and Dafont is used for downloading additional fonts. For the pre-tasks, they would go through the websites and look at how cluttered they were before installing the extension. They were not required to find or click on anything but they were able to explore the website if they wanted to. Once they were content and had read through and looked at both websites, they were asked to close the tab. The final pre-task contained a link to the extension on the Chrome web store. After they downloaded it, it was automatically enabled and they could continue with the tasks.

5.1.2 Tasks

The first task that was created took them to Softpedia with the extension enabled. The link that is given in the task takes them to the information and download page of the program OnyX. With the extension installed now, the advertisements have been highlighted or removed. They were asked to find the real download button for the program and were asked if they were able to find it. They were not required to specify where it was or download anything, just say if they found it or not. The website for this task was found from Brignull's list of examples and was one of the top examples for Disguised Advertisements [3]. The goal

for this task is for the user to be able to navigate a website successfully with the extension being active and highlighting and removing the advertisements. By seeing the differences from the pre-task, we are hoping to have them be more aware of disguised ads. To be considered a successful completion, the participant had to find the real download button and report it.

The second task was similar to the first and had the participants go to DaFont and find the download button for a font. Similar to the previous task, the advertisements have either been highlighted or removed. Unlike task one, the participants were required to give a short answer if they were able to find the button and where it was on the webpage. The website for this task was found by the researchers when creating the tasks. It was chosen due to its small download button that is hard to notice due to being positioned away from the main content of the page (on the middle-upper right side) and having a non-bright color. The goal of this task was similar to Task 1 in seeing the disguised ads but also was to show the users the different kinds of websites that have dark patterns and how they vary in how they are shown. Similar to task 1, a success was defined by if the participant could find the position of the real download button.

The third task had the participants go to CNN's business webpage. On this page, there is a list of different articles which also contained an advertisement article [4]. The extension would highlight the article title and the participants were asked to enter the title of the article. Due to the website changing articles over time, the answer given was expected to vary across participants depending on what day they completed the survey. This task came from researchers exploring websites and looking for tasks. It was selected because it had advertisements that were made to look like an article and were added to a list on the website. The goal of this task was to have users be able to see a different way that disguised ads and

dark patterns can be hidden and that they are used on popular websites as well. A success for this task was defined by if the user could give the name of the correct advertisement that was highlighted in the article list.

The fourth task was done on the website GoDaddy which is commonly used to buy domain names. The participant was asked to simulate the process of buying a .net website without buying it. They were explicitly told that they were not required to purchase anything to complete the tasks and overall survey and that they just needed to continue to the cart page. They were asked to check the price that was displayed at the beginning of the task and to enter what the final price was as a check to make sure they went through the process and saw the price increase. Finally, they were then asked if a notification appeared warning them about a potential price change. The website for this task was found on Brignull's website as one of the examples for Sneak into Basket [3]. The goal of this task was to show users what Sneak into Basket is and how websites use it and how the Dark Pattern Detector alerts them at the cart. It was intended to show how easily websites are able to add things to the basket, how many times they try, and how they try and convince and persuade the user into thinking they are getting a good deal. This task was considered a success if the notification popped up. Entering in the correct price was a check to see if they were properly following instructions.

The fifth and final task was on the website ProFlowers. The participants were asked to simulate purchasing a specific bouquet of flowers and keep track of the original price. For delivery information, they were told to enter the address of McBryde Hall on Virginia Tech's campus to keep prices consistent. After completing the process and entering the delivery information, they were asked to enter the final price of the arrangement. Participants were told to keep the default options when completing the task for consistency between users.

This website was also found on Brignull's website as an example for Hidden Fees [3]. The goal of this task was to have users experience the process that websites have users go through before adding the fees to the total and having them see the notifications from the extension. This task shared the same success condition as task 4 of having the notification pop up and entering the final price to make sure they went through the process properly.

5.1.3 Exit Questions

Once the tasks were completed, the participants were given instructions on how to remove the extension. They were asked if they were able to successfully remove it and if there were any difficulties in doing so. Then, the participants were given a series of questions about the extension.

The first questions asked what the participants' opinions on the extensions were overall, what specific parts they liked or disliked, and what they would change about the extension if anything. We wanted to get specifics on what the users thought about the extension and what was popular and what was not. Next, we asked the participants about the visuals of pattern detection. First, we asked if the color from the highlighting was over the top or too much/too invasive and then asked the same thing about the notifications. Following them, we also asked for any additional notes that they would like to add. These were to get feedback on our design choices and see how the users responded to them. The next question asked if there were any additional dark patterns or deceptive designs that they would like the extension to cover to see where we could look for future works and to build on the extension. Finally, we asked if there was anything else that they wanted to mention that happened as. This was a way to give the participants a chance to add any comments about bugs or ask questions that we may not have given a previous space to mention.

Chapter 6

Results

Overall, the extension was given mostly positive reviews and was found to be effective for completing the given tasks. Out of the 40 participants, 29 gave entirely positive feedback saying that they liked it or claimed that they thought it was good at detecting dark patterns and for helping them learn about what they need to do or not do involving dark patterns. In the remaining 11, six gave feedback about disliking the yellow highlighting for ads with five of them still giving some positive feedback, two gave a mix of likes and dislikes about the extension and gave critiques on the notifications system, and the final three gave only complaints or had issues with the extension's functionality.

6.1 Task Results

At the start of the survey, we asked participants about their familiarity with dark patterns and browser extensions. Out of the 40 participants, 20 claimed that they were unaware of what dark patterns were. Of the remaining 20 participants, six of them claimed to be only semi-familiar with dark patterns with the remaining 16 saying that they knew what they were. For the extensions, only nine of the participants said that they had never used browser extensions with an additional eight claiming that they have used browser extensions

intermittently in the past. There did not appear to be any correlation between those that knew what dark patterns were and those that used extensions.

Each of the tasks' results was made into a percentage showing the rate at which the tasks were completed successfully. For the first task, 87.5% (35/40) of participants were able to locate the download button properly with the extension enabled. The second task had 82.5% (33/40) of participants giving the correct answer and finding the download button on DaFont. However, of the seven participants that did not give the correct answer, only three claimed that they could not find the download button at all. The remaining four participants either gave a vague answer saying that they found it but not specifying where it was or simply answered with the phrase "Done". Since they did not give a proper answer, we counted them as incorrect which reduces the success rate by 10% (92.5% to 82.5%). For the third task, 95% (38/40) were able to correctly identify the article in the list. The two incorrect responses for this task were also the phrase "Done" from the same two participants.

The fourth task contained the lowest success rate with a 72.5% (29/40) that had the notification appeared for them on the checkout page of GoDaddy. This varied from the fifth task which had an 82.5% (33/40) success rate for the notification appearing for ProFlowers. Of the participants the notification did not appear for, only one individual did not have the notification appear for the fifth task but did have it appear for the fourth.

Task #	Success Rate
1	87.5%
2	82.5%
3	95%
4	72.5%
5	82.5%

Table 6.1: The success rate of each individual task.

6.2 Exit Survey Results

The participants' answers from the exit survey were collected and were put through Open Coding Process, a type of qualitative analysis. Open Coding Process is done by breaking qualitative data into single words or short expressions that are meaningful and describing them [9]. This is done by two different researchers separately to come up with themes and descriptions and see if what they determine matches. The main themes that we found for our results were broken down to whether or not they liked the extension and the reasons behind why they did or did not.

For the first theme, it was found that 35 of the 40 participants liked the extension. Of the remaining five, 4 disliked it with 1 participant's answers not giving a clear enough side and was placed in the middle. The most popular reasons for liking the extension were "identifies dark patterns" with seven participants. The next most common was that it "raises awareness" with five participants and "useful" with four. Other reasons for why with fewer participants were "saves time", "real-time detection", "pop-up alert", "clear display", "works as expected", and "interesting". For dislikes, each of the four participants had different reasons. First, one claimed the "yellow was out of date" that was used for highlighting the advertisements. The next two were "Notifications did not properly appear" and "[the extension] Did not work well for them". Finally, the last participant's reason was "Did not like the tasks" which is not specifically a dislike against extension but was still their main reason. The remaining seven participants, including the 1 that was placed in the middle, did not give enough details in the answer to determine a why.

Chapter 7

Discussion

In this study, we found that the majority of the users were able to successfully complete the tasks and found the extension that we made to be useful. Some of the positive comments were "It was a really nice idea. It will definitely save time when we are trying to find the right download button" (Participant 8), "The extension helped highlight deceitful patterns on a website and made my life a lot easier" (Participant 28), and "It helped in understanding the dark patterns and helped in understanding [to] not click on links unnecessarily" (Participant 33). However, some participants did not find it useful or had problems running it. Participant 11 claimed that it did not help them find the download buttons in the first and second tasks and wished that the extension was more accurate. Additionally, as mentioned at the beginning of the results, 15 of the participants complained that they thought the yellow was too bright or invasive. Most of the remaining 25 had either no opinion or said that they found the yellow color to be a fitting choice. Participants 26 and 31 stated that they thought the notifications and highlighting could have been more invasive and noticeable in order to guarantee that the extension's users did not accidentally miss a notification or were unaware of what was an advertisement.

7.1 User Feedback

The users gave a lot of feedback on what they would like to see in future versions of the extension. When asked if there were any other dark patterns that they would like the extension to cover, 14 participants gave responses both for new patterns and additions to the patterns that are currently covered. The responses were split among the pattern types of Forced Continuity which is when a user is silently charged after a free trial runs out, Roach Motel which is when it is difficult to leave or unsubscribe from a service, and Confirmshaming like in Figure 2.5. In addition to these patterns, they also gave suggestions on features that either related to the current extensions or ideas that did not fall under a specific category. Two participants suggested un-checking boxes that are automatically selected such as "Select this box to subscribe to our newsletter" boxes or like the one that was selected by GoDaddy seen in 2.3 and another two asked for more information about the dark patterns that the extension detects. Other suggestions included having a message appear when a user hovered over an ad letting them know what it was and highlighting buttons that are hard to see.

7.2 Limitations and Future Work

While creating the extension and survey and having participants complete the study, there were limitations and issues that occurred. Initially, we spent time deciding what patterns we wanted to detect, how we wanted to display and highlight the patterns, and how we wanted to distribute our program. We decided on creating an extension and publishing it as it would allow the participants to easily install and use it. However, it required us to set up and pay a small fee to create a web developer account. In order to publish our extension, we had to have it approved by the web store and any permissions that we used would add time

and potential issues for getting it approved. The approval process took time which made it harder to make additional changes if any issues came up since we wanted to get the survey out and have data collection completed in a timely manner. For the dark patterns and how we displayed them, we were limited by the types of dark patterns we could complete a study on. To complete the study in a timely manner, we had to limit the types of patterns and the number of patterns that were added to the extension.

For displaying the disguised ads, we had to decide which color would work best for highlighting on the web. Since many websites contain black or darker-colored text, darker colors were not able to be used as it made the text unreadable. If we used these colors, advertisements that are links to articles, such as the ads on CNN's business page, the user might still be interested in being able to see what it is and potentially click on it. Most pale or lighter colors were also unable to be used as they were not as noticeable and could be overlooked by the user. While the yellow we choose fell in the previously mentioned restrictions, tritanopia, a type of yellow-blue colorblindness, causes yellows to appear as almost white which would make it look like the background of the webpage [14] [5]. To combat this, we added a black border around the ads to still make them visible to anyone who is affected by this. In the future, we would like to either find a highlighting method that is not affected by any type of colorblindness or add colorblind modes that can be toggled to change the highlighter color based on what type of colorblindness the user has. For the notification, we were limited by websites being inconsistent in how they represented the total price of what was being purchased. The differences made it difficult to track the price consistently across different websites and across different pages on the same website.

An important limitation that came from the extension was the appearance of false positives. False positives are when items that did not necessarily fit the criteria of a dark pattern were detected (specifically, a dark pattern that we were trying to detect). An example of this can

be seen in the image shown from Softpedia from task 1. The ”ad” detected is a donation link for the creators of the software. While this is not necessarily an advertisement, it does fall on the fine line of still being a dark pattern, specifically donationshaming. This is similar to confirmshaming in how it is trying to guilt the user into doing something, in this case convincing the user to donate. While this example is a false positive because we were not looking for confirm or donation shaming currently, it would be hard to distinguish if this is a dark pattern or not automatically. Additionally, other dark patterns could also have false positives and we are unable to tell how many times this occurs in our extension.

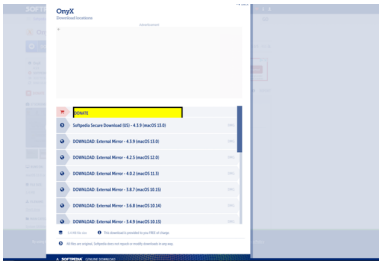

Figure 7.1: The download page on Softpedia with the extension turned on. The highlighted area says ”Donate” and links to the creators of the software.

For continuing work on the Dark Pattern Detector, we would need to reduce and eventually remove all instances of false positives. The main method for this would be to go through and find different instances of false positives and look for the causes in the extension’s logic. This would allow us to see if there needs to be small fixes or a complete overall of the program’s detection logic to remove the errors. For additional patterns like confirmshaming or any type that requires an understanding of what the words mean, we would likely require a manual choice from the developers if it should be considered a dark pattern.

Another change for future studies comes from changing the appearance of how the dark patterns are shown to the user. For the notification system, some study participants requested a larger pop-up to appear. In future studies, we would add different sized, colored,

and positioned notifications to understand what most users preferred and what was most effective in informing them. For the color of the highlighting of patterns, future work is broken down into two different parts. First, similarly to the notifications, testing different highlighting colors for user preference to see if yellow is the most preferred option or if there are other colors that the users like. Next, we also want to add color blindness switches to make the extension more accessible. A study done with color blind participants to find which colors are most visible for each type would be beneficial for our extension and other detection programs.

We also look to add additional dark patterns for the extension to detect. The most common three user responses were Confirmshaming, Roach Motel, and Preselection. For preselection, as seen in the Godaddy example in Figure 3, we want to add in a way of alerting the user that a box has been initially selected instead of automatically selecting a no/deny option. This would be to prevent the extension from also preselecting something for the user. The notification could either appear where the check box that was selected is or like the notification system used for the Sneak into Basket and Hidden Fees used by Dark Pattern Detector. Roach Motel would be similar to Disguised Ads by highlighting where unsubscribe buttons are or paths to which the user could leave the subscription. Finally, confirmshaming would need something such as natural language processing to determine if the options given are trying to guilt and convince the user to select a specific option. To do this, examples of confirmshaming should be used in the training and then studies done across a wide range of websites that may or may not have confirmshaming.

Alongside adding more dark patterns, Artificial Intelligence and Machine Learning need to be looked at both in helping the extension detect dark patterns but also in how they may lead to websites being deceptive. In a Harvard Journel of Law & Technology volume titled ”Deception by Design” by Lauren Willis, she discusses how businesses use different

tools to automatically create advertisements, processes, products, and services in a way to optimize profits [19]. In their optimizations, they may end up unintentionally (or intentionally) adding dark patterns in if they are the most effective way to generate clicks or increase revenue [19]. With these methods, it can become difficult to determine whether a deceptive design is intentional or not as the border between what is a dark pattern and what isn't is blurred and to detect them as they get hidden more effectively. A way to combat this could be teach a program, such as an artificial intelligence, to be able to detect dark patterns in order to keep up with changing technologies and practices in business. However, it would encounter a similar roadblock that humans have which is determining the intent of design. For example, for detecting confirmshaming, the artificial intelligence would have to be able to properly determine the difference between a service that is trying to guilt the user and a service that is being advertised to them without any malicious intent. This could lead to false positives or missing out on some cases entirely depending on the training for the tool.

Finally, we also want to make the extension more accessible and give more information about dark patterns. The first step is to have the extension be available on browsers outside of Google Chrome. The process for creating an extension for other browsers is similar to Chrome but requires developer accounts for their store fronts and differently formatted manifest and options files as well as approval processes. The JavaScript that is used for the extension is compatible with most modern browsers but each one would need to be checked individually and adjusted accordingly.

To help give the users understand dark patterns and how the extension works, we suggest three different ideas. The first is to create an introductory tutorial for using the extension.

While the pre-tasks were useful in introducing what the websites would look like and how they would change with the extension, they did not explain what was happening. For future versions, we want to add examples with explanations for each type of dark pattern that is being detected. The last two ideas are for patterns that we already detect. For disguised ads, we look to add a message that would appear when hovering over an advertisement as an accessibility feature and a way to give the user another warning if they might not have noticed that it was highlighted. For the notification system, we want to add a breakdown of what changed to the notification by saying what item was added, what fees were added to the cost, and the difference in cost from before and after these were added to the total.

After changes have been made, additional studies should be done both for testing the extension itself and for helping teach the participants more about dark patterns. For the extension, our study used a set of pre-selected websites that we knew had dark patterns on them. A study that lets participants tests the extension by using it in their daily lives and mark what patterns were detected on what sites. This would allow for the researchers to get a better grasp of what kind of false positives can appear and how accurate the extension is. Additionally, by using it for a long time, participants could give thorough, more detailed feedback about the extension. This kind of study would need to be evaluated differently than this study that used tasks and a survey as a way to understand what people thought of the extension and if it worked for them. To evaluate the teaching aspect of the extension, participants can be asked about dark patterns before and after using the extension to see how well they comprehend them and the difference that the extension made. This could be done after using the extension short-term and then asking participants after a longer period of time to see how well it retained.

7.2.1 Study Questions

In the user study, there were a few interesting cases that we found in our data. First, for the fourth and fifth tasks, we found that users would click on different options outside of what was asked. For example, they would select different options causing slight fluctuations in price. The main purpose of these tasks was to see if the extension worked and if the notification appeared and if the changes did not prevent it from appearing. Because of these differences, we want future studies to require a specific formatted answer and process to prevent the additional time required to go through and make sure the changes did not cause any issues.

There were also three separate cases of issues with some of the participants' answers. The first case was that two participants shared identical names. The answers that they gave were different and their responses were submitted nine days apart so we kept both responses in the data. For future studies, we will collect additional identifiers in order to increase confidence that there were no duplicate submissions. The second case had two participants submit similar answers with a small time window between their survey submissions. This case had both participants be the users who submitted the answer "Done" to the third and fourth tasks which gave the impression that they were completing the survey together. Since the other answers and demographic information was different, we kept the data in order to go collect more data on the effectiveness of the extension but did not count the "Done" answers as correct. The last case was for Participant 23 who mentioned in his feedback that he had the notification box appear multiple times on the ProFlowers website and that DaFont did not contain any ads before the extension was installed. Additionally, he mentioned that he completed the survey on Ubuntu Linux using Chromium and was the only participant to mention either issue. We were unsure if his setup caused any of these issues and in the future will ask for additional information about each participant's system to see if the

different setups cause any bugs.

Chapter 8

Conclusions

For this study, we wanted to look into dark patterns on the web and how we could help individuals avoid them and learn more about them. To do so, we investigated how familiar people were with dark patterns, how comfortable people are with downloading and using browser extensions, what ways can users navigate websites with dark patterns, and what people thought of the Dark Pattern Detector. To do this, we had them use the Dark Pattern Detector and complete a series of tasks and answer questions to see what they thought of the extension and dark patterns.

We found that half of the 40 participants had no familiarity or prior knowledge of dark patterns and an additional six were only semi-familiar with them. We also saw that 31 of the 40 participants had used browser extensions making it a potentially acceptable method for spreading information about dark patterns. Additional studies that include other distribution methods such as an app version to compare it against are recommended. Through Open Coding Process we found that 35 of the 40 participants were found to like the extension, with four of the remaining five disliking it and the last participant being neutral. For the functionality, the majority of the participants successfully completed the tasks with the success rates being 87.5% for Task 1, 82.5% for Task 2, 95.0% for Task 3, 72.5% for Task 4, and 82.5% for Task 5. These results showed that most were able to successfully navigate through

the websites with dark patterns. While some of the percentages may have been brought down by potential misunderstandings from the participant or issues from their setup, that means that for future studies, the questions can be made clearer to make sure everyone has a fair chance to understand it and have clearer system requirements. Additionally, we also want to make the extension work on any type of system so anyone can use it and it is not limited to just one group. With the positive response from the participants, we hope to expand upon this in the future by continuing the use of extensions and through creating a baseline for furture works to add additional patterns to.

Bibliography

[1] 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 *Designing Interactive Systems Conference 2021*, pages 763–776, 2021.

[2] Christoph Bösch, Benjamin Erb, Frank Kargl, Henning Kopp, and Stefan Pfattheicher. Tales from the dark side: privacy dark strategies and privacy dark patterns. *Proc. Priv. Enhancing Technol.*, 2016(4):237–254, 2016.

[3] Harry Brignull. Decptive patterns. URL deceptive.design.

[4] CNN. Cnn business. URL https://www.cnn.com/business.

[5] Colblindor. Trianopia - blue-yellow color blindness. URL https://www.color-blindness.com/tritanopia-blue-yellow-color-blindness/.

[6] Gregory Conti and Edward Sobiesk. Malicious interface design: exploiting the user. In *Proceedings of the 19th international conference on World wide web*, pages 271–280, 2010.

[7] Andrea Curley, Dympna O'Sullivan, Damian Gordon, Brendan Tierney, and Ioannis Stavrakakis. The design of a framework for the detection of web-based dark patterns. 2021.

[8] Linda Di Geronimo, Larissa Braz, Enrico Fregnan, Fabio Palomba, and Alberto Bacchelli. Ui dark patterns and where to find them: a study on mobile applications and user

perception. In *Proceedings of the 2020 CHI conference on human factors in computing systems*, pages 1–14, 2020.

[9] Uwe Flick. *An Introduction to Qualitative Research*. SAGE, 2009.

[10] GoDaddy. Godaddy domain names. URL godaddy.com.

[11] Philip Hausner and Michael Gertz. Dark patterns in the interaction with cookie banners. *arXiv preprint arXiv:2103.14956*, 2021.

[12] Arunesh Mathur, Gunes Acar, Michael J Friedman, Eli Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan. Dark patterns at scale: Findings from a crawl of 11k shopping websites. *Proceedings of the ACM on Human-Computer Interaction*, 3 (CSCW):1–32, 2019.

[13] 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*, pages 1–18, 2021.

[14] Medline Plus. Color vision deficiency. URL https://medlineplus.gov/genetics/condition/color-vision-deficiency/#frequency.

[15] S Hrushikesava Raju, Saiyed Faiayaz Waris, S Adinarayna, Vijaya Chandra Jadala, and G Subba Rao. Smart dark pattern detection: Making aware of misleading patterns through the intended app. In *Sentimental Analysis and Deep Learning: Proceedings of ICSADL 2021*, pages 933–947. Springer, 2022.

[16] Daniel Ruby. 67 gmail statistics for 2023. URL https://www.demandsage.com/gmail-statistics/.

[17] Than Htut Soe, Oda Elise Nordberg, Frode Guribye, and Marija Slavkovik. Circumvention by design-dark patterns in cookie consent for online news outlets. In *Proceedings of the 11th nordic conference on human-computer interaction: Shaping experiences, shaping society*, pages 1–12, 2020.

[18] Than Htut Soe, Cristiana Teixeira Santos, and Marija Slavkovik. Automated detection of dark patterns in cookie banners: how to do it poorly and why it is hard to do it any other way. *arXiv preprint arXiv:2204.11836*, 2022.

[19] Lauren E Willis. Deception by design. *Harv. JL & Tech.*, 34:115, 2020.

Appendices

Appendix A

First Appendix

A.1 Section one

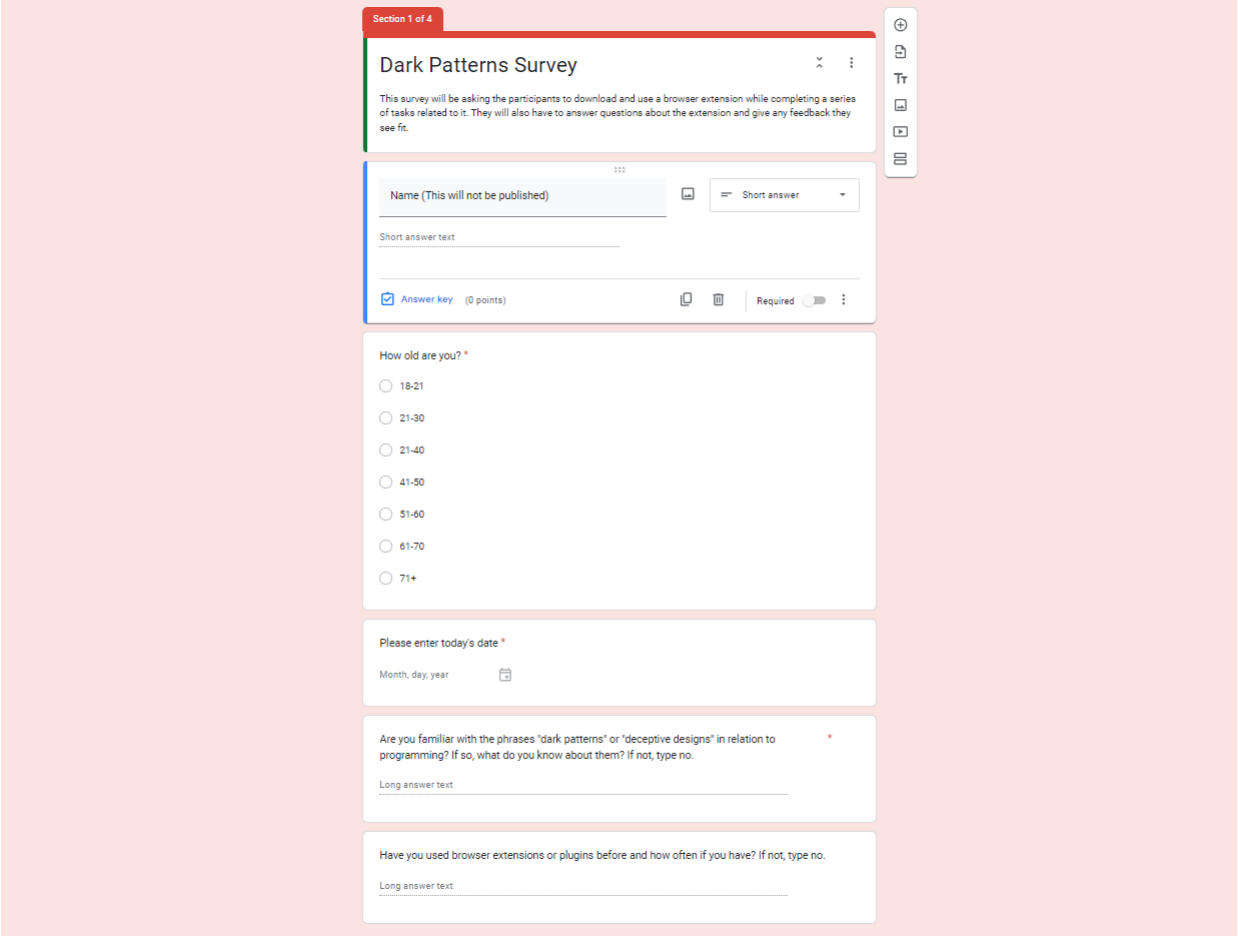

Figure A.1: The first section of the survey the participants had to complete. This contained the demographics questions and the dark pattern and extensions questions.

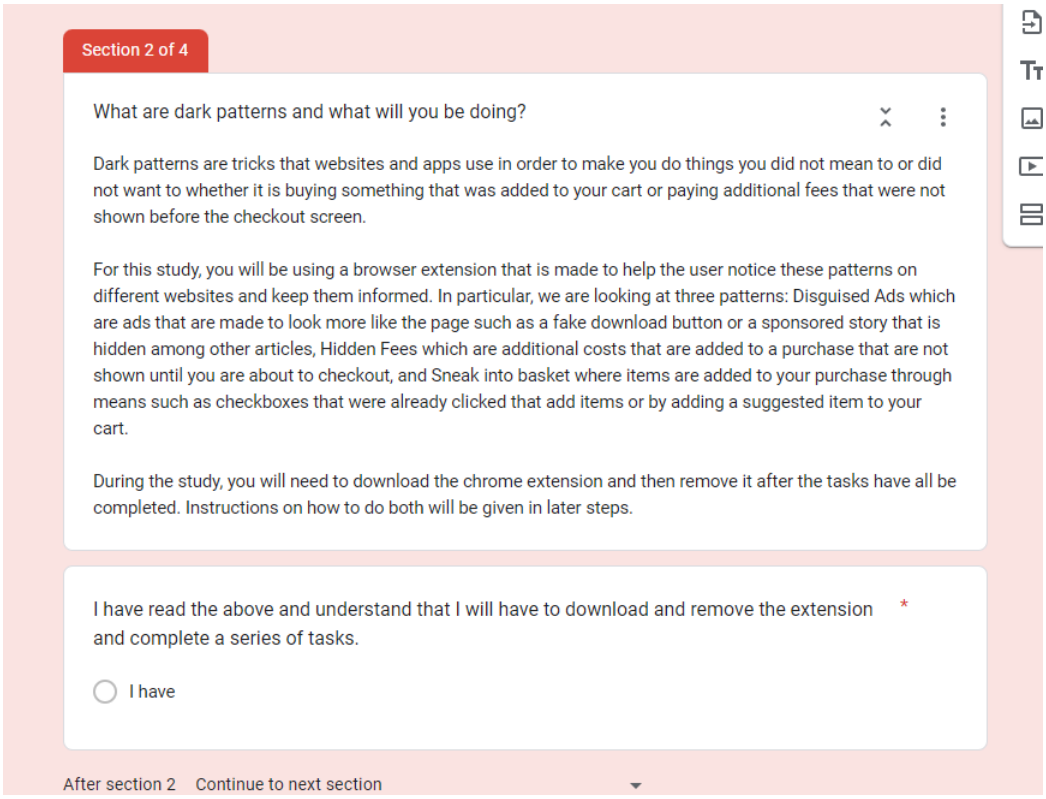

Figure A.2: The second section of the survey. This contained information about what the participant would be doing and information about dark patterns.

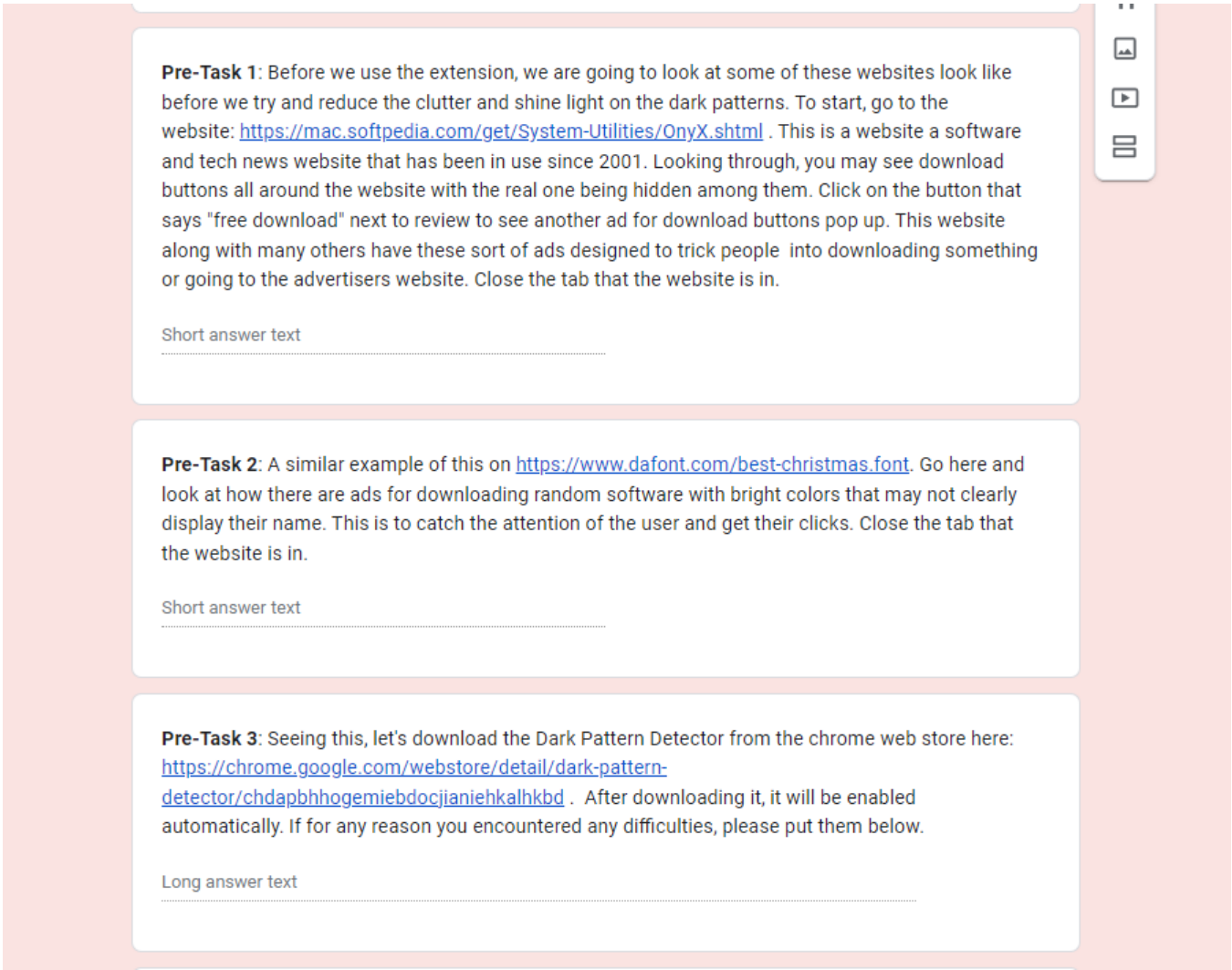

Figure A.3: The beginning of the third section of the survey which contained the pre-tasks that the participant had to complete.

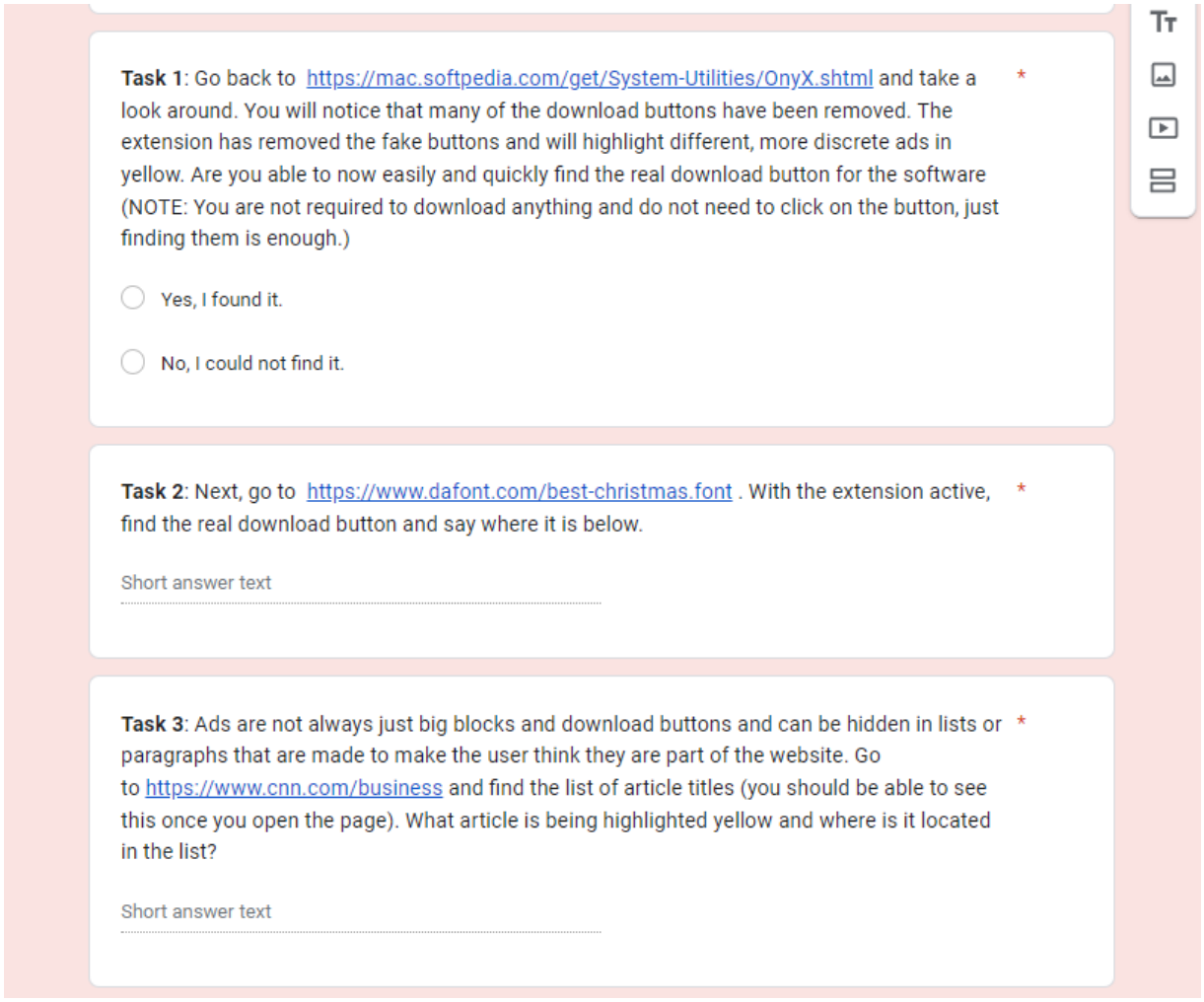

Figure A.4: The first three tasks in the survey.

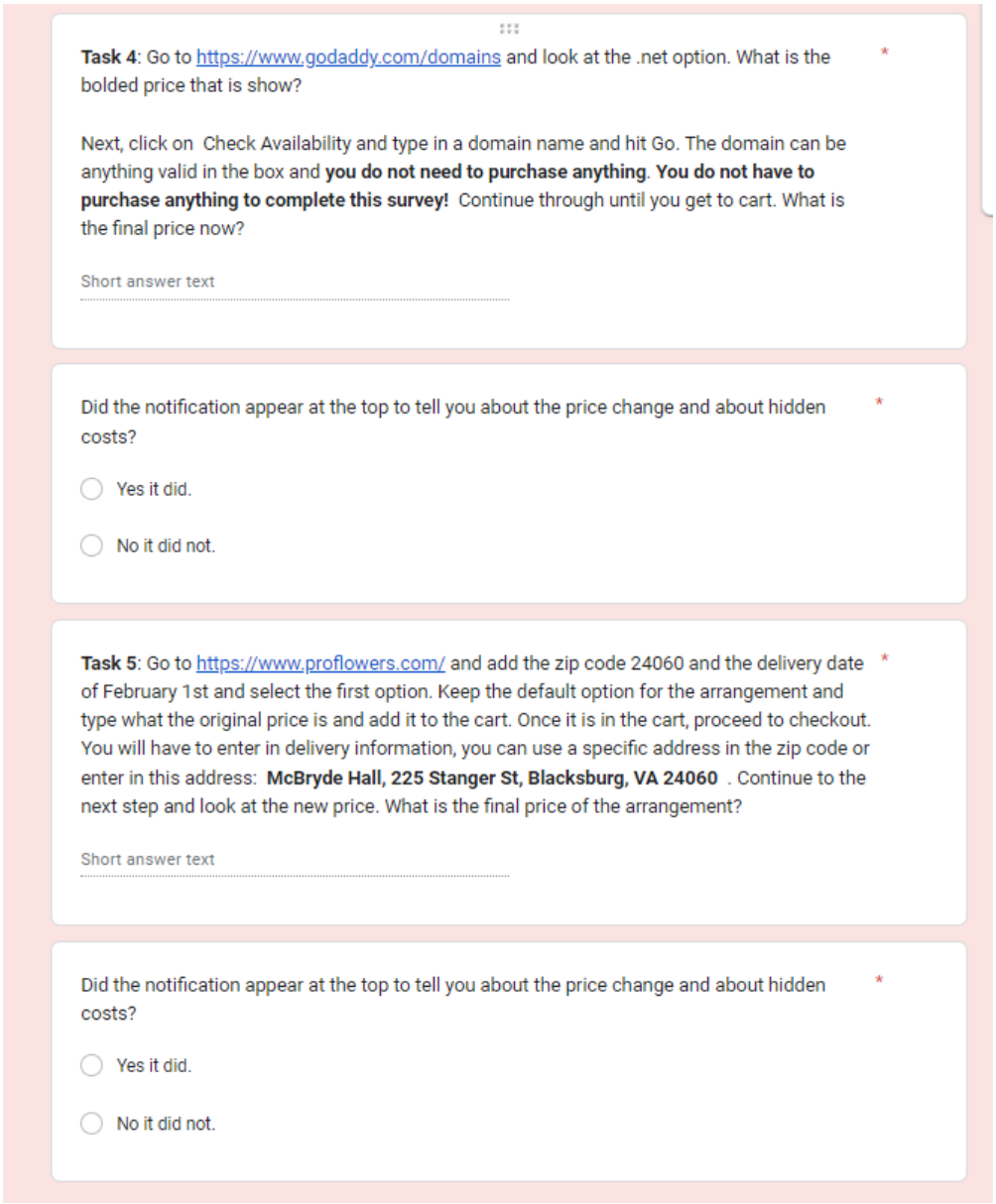

Figure A.5: The fourth and fifth tasks in the survey.

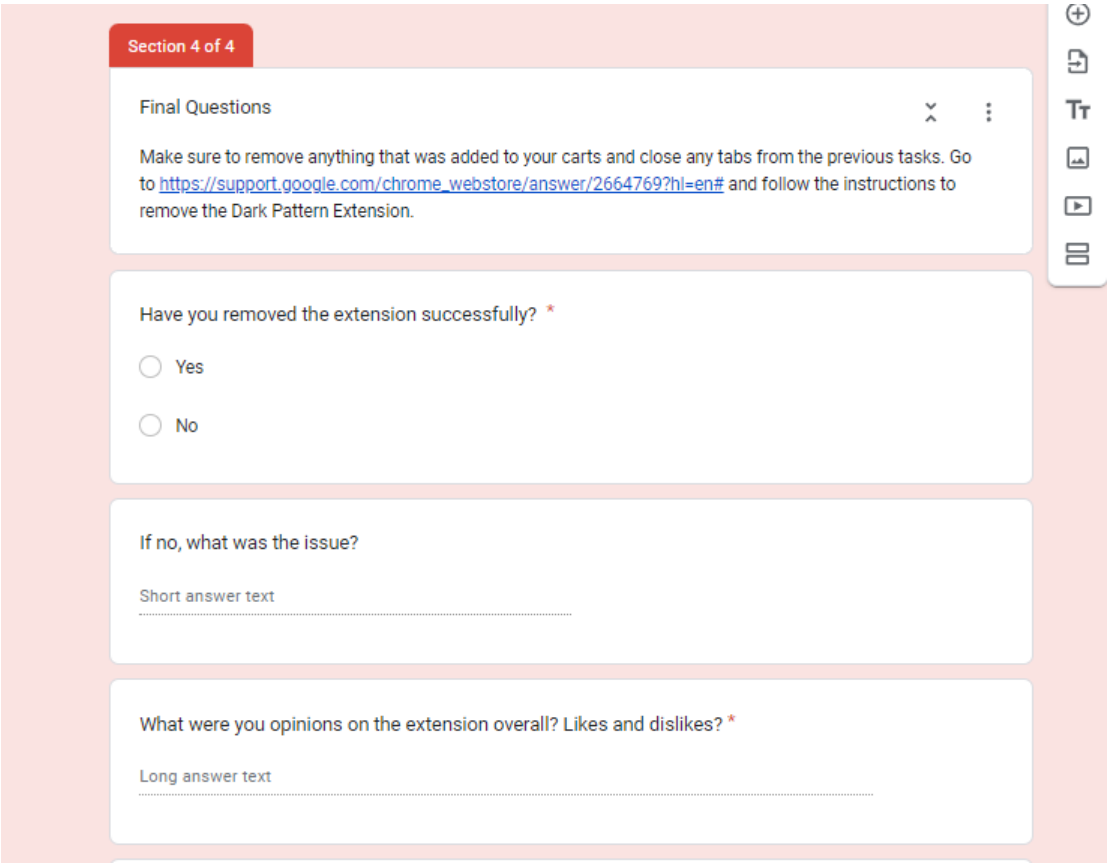

Figure A.6: The first half of the post-survey questions for collecting feedback.

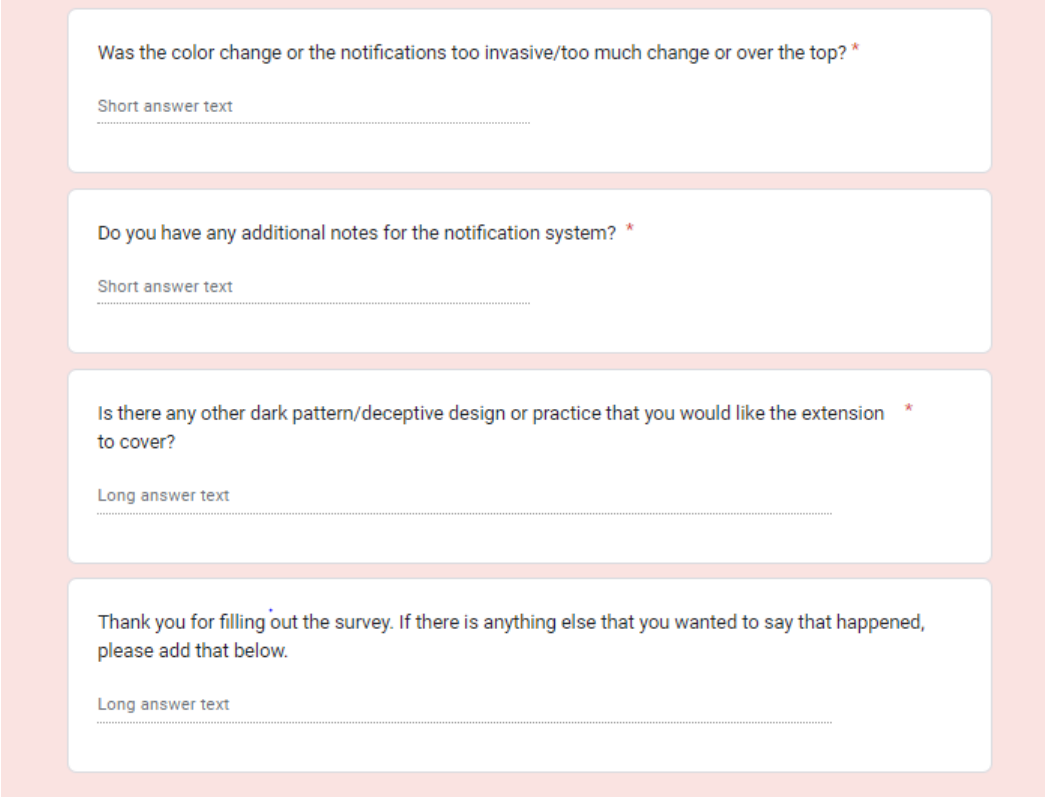

Figure A.7: The second half of the post-survey questions for collecting feedback.